Italia '90
Il primo passo della HDTV digitale
II parte

Serie di articoli, pubblicati in più numeri di Elettronica e Telecomunicazioni, trattano e approfondiscono una singola tematica. Lo scopo dell'iniziativa LeMiniSerie è di raccogliere tali articoli, con una veste tipografica unitaria che ne faciliti la consultazione e apportando correzioni e aggiornamenti ritenuti opportuni.

Questa è la seconda parte del terzo volume.

Italia '90 - Il primo passo della HDTV digitale raccoglie 20 articoli pubblicati su Elettronica e Telecomunicazioni negli anni 88-93. Tali articoli illustrano le tecniche e le realizzazioni che portarono al successo del progetto di diffusione via satellite, per la prima volta in assoluto, di immagini televisive in alta definizione digitale.

LeMiniSerie sono una iniziativa del
Centro Ricerche e Innovazione Tecnologica della
www.crit.rai.it

In copertina:
Apparati per la decodifica dei segnali HDTV digitali
alla sede Rai di Venezia (Palazzo Labia)
in occasione di Italia '90.

Italia ‘90 Il primo passo della HDTV digitale - Parte II

Vent’anni fa, per la prima volta, un segnale HDTV digitale fu diffuso via satellite. Era la prima partita del Campionato Mondiale di Calcio Italia ‘90.

Tale evento era stato possibile grazie ai risultati del progetto europeo Eureka 256 proposto e attuato da un consorzio italo-spagnolo costituito da Rai e Telettra, per quanto riguarda la componente italiana, e da Telettra Española, Retevision e il Politecnico di Madrid, per la componente spagnola.

L’evento è ricordato nella terza raccolta de LeMiniSerie, per facilitare la consultazione la pubblicazione è divisa in due parti, la prima parte è introduttiva, mentre questa seconda parte raccoglie i 20 articoli pubblicati da Elettronica e Telecomunicazioni su tale progetto a partire dal 1988 fino al 1993.

Torino, dicembre 2010

Indice degli Articoli

Indice degli Autori

CODIFICA DI SEGNALI TELEVISIVI AD ALTA DEFINIZIONE PER LA TRASMISSIONE NUMERICA VIA SATELLITE

G. BARBIERI *

SOMMARIO — I tempi tecnici necessari per l'introduzione di un vero e proprio servizio di televisione ad alta definizione (HDTV) dipendono in larga misura dalla disponibilità di canali di trasmissione ad elevata capacità, per applicazioni che riguardano sia il contributo che la distribuzione. Lo sviluppo tecnologico in atto lascia prevedere che in futuro verranno sviluppate reti di collegamento operanti ad una velocità di trasmissione di 565 Mbit/s; tuttavia, l'esigenza di avviare in tempi brevi alcune sperimentazioni tecniche di trasmissione impone di far fronte ad un certo numero di problemi legati soprattutto alla capacità relativamente contenuta degli attuali canali. Le tecniche di riduzione del bit rate sono oggi oggetto di studio in molti laboratori. Nel presente lavoro vengono evidenziati gli aspetti tecnici associati alle strategie di codifica di sorgente applicabili alla HDTV; inoltre, vengono discussi i risultati di una serie di indagini svolte sull'algoritmo DCT, mediante simulazione all'elaboratore elettronico, aventi lo scopo di valutare i fattori di merito e gli eventuali limiti del suddetto approccio. Sulla base dei risultati finora conseguiti, sembra possibile sviluppare un sistema di trasmissione HDTV numerico operante ad un bit rate lordo inferiore a 140 Mbit/s; sono in corso studi per verificare se, a più lungo termine, sia altresì realizzabile un servizio di diffusione via satellite nella gamma dei 60÷70 Mbit/s.

SUMMARY — ***Digital encoding of HDTV signals for applications in satellite transmission.*** *The time scale for the introduction of a real HDTV service, to a great extent depends on the availability of high-capacity transmission channels, both for contribution and distribution applications. The current technological advance makes it possible to foresee that digital networks, operating at a bit-rate up to 565 Mbit/s, will be developed in the future; however, in order to fulfill the requirements for a short term starting of some broadcasting experiments, the operational constraints posed by the bandwidth limitation of the existing channels must be taken into account. Bit-rate reduction techniques are currently being investigated in many laboratories. This paper outlines the technical aspects relevant to the source encoding strategies applicable to HDTV; moreover, it discusses the results of a campaign of investigations carried out on DCT algorithm, by means of a computer simulation, with the aim of evaluating the factors of merit and the limits of the above approach. On the basis of such results, the development of a HDTV digital-broadcasting system operating at a gross bit-rate lower than 140 Mbit/s appears to be possible; studies are in progress in order to ascertain if, at longer term, a satellite broadcasting in the range of 60-70 Mbit/s is also feasible.*

1. Introduzione

I primi tentativi di sperimentazione sulla televisione ad alta definizione risalgono al 1968 ma è solo con l'avvento degli anni '80 che gli studi per la formulazione delle specifiche di sistema e lo sviluppo dei relativi apparati hanno subito un impulso vertiginoso.

In Giappone è stato varato un piano per l'avvio, entro l'inizio degli anni '90, di un vero e proprio servizio di radiodiffusione via satellite di programmi HDTV (High Definition TV).

In Europa, a parte i vari esperimenti di produzione elettronica di film compiuti finora dalla RAI, esiste un progetto interindustriale, supportato dalla Comunità Europea (progetto Eureka EU95), avente come scopo lo sviluppo di una linea di apparati HDTV che dovrebbero coprire una vasta gamma di applicazioni sia nel settore della produzione dei programmi, sia in quello della ricezione. Le odierne realizzazioni nel campo delle fibre ottiche e dei satelliti, nonché i recenti progressi nelle tecnologie VLSI (Very Large Scale Integration), contribuiscono a rendere fatti in un futuro relativamente prossimo, una catena HDTV completa, dalla sorgente al ricevitore domestico.

L'uso di mezzi di trasmissione numerica, per reti sia di contributo che di distribuzione, impone che vengano risolti i problemi derivanti dall'elevata velocità di trasmissione richiesta dagli standard di produzione attualmente proposti.

Metodi per la riduzione del bit-rate sono oggetto di studio in vari laboratori; gli obiettivi perseguiti sono focalizzati su una vasta gamma di applicazioni, dalla videotelefonia alla televisione ad alta definizione. Molti di questi studi sono basati su principi comuni ma le diverse applicazioni impongono la scelta delle tecniche ritenute di volta in volta più appropriate.

Presso il Centro Ricerche della RAI è in corso una intensa attività di ricerca, condotta in collaborazione con l'industria nazionale, sui metodi di codifica numerica per segnali televisivi; la sperimentazione finora effettuata ha avuto come oggetto il segnale confor-

---

* Ing. Gianfranco Barbieri del Centro Ricerche RAI-Torino.
Dattiloscritto pervenuto alla redazione il giugno 1988.

me alla Raccomandazione 601 del CCIR (Comitato Consultivo Internazionale di Radiodiffusione) ed ha portato a contribuire alla attività normativa di alcuni organismi internazionali (CCIR, EBU) (bibl. 1, 2).

Gli studi attualmente in corso utilizzano come punto di partenza i suddetti risultati ed hanno come obiettivo il segnale televisivo numerico ad alta definizione.

2. Metodi di riduzione del bit-rate sviluppabili con le attuali tecnologie

Le odierne tecnologie offrono una vasta casistica di metodi per la riduzione del bit-rate di segnali televisivi; una descrizione esauriente di tali metodi è riportata in bibl. 3. In sintesi, si possono fare le considerazioni seguenti:

- In alcuni sistemi (ad esempio, quelli basati su tecniche DPCM/Differential Pulse Code Modulation) il segnale è trattato nello stesso dominio spazio- temporale nel quale è stato generato. In altri, esso è trasformato in un dominio di tipo «spettrale» prima di essere sottoposto a riduzione di ridondanza.
- Alcuni metodi fanno ricorso ad un sottocampionamento fisso (generalmente bi o tri-dimensionale) a cui è inevitabilmente associata una azione di filtraggio; in altri, più sofisticati, la rimozione di componenti spettrali avviene in modo adattativo in funzione del contenuto di informazione dell'immagine.
- I sistemi più semplici trattano l'informazione associata al solo contenuto spaziale mentre quelli più complessi agiscono anche sulle proprietà temporali, dovendo ricorrere, pertanto, all'uso di memorie di quadro.
- Nei sistemi più sofisticati si utilizzano tecniche di rivelazione e compensazione del movimento.

La scelta dell'algoritmo ottimale è frutto di un compromesso tra la complessità del codec e il costo della trasmissione; in particolare, la complessità del codec è un argomento difficile da puntualizzare data l'elevata dinamica di sviluppo dei circuiti VLSI.

Gli algoritmi basati su tecniche predittive possono assumere forme relativamente semplici se limitati al dominio spaziale associati a quantizzatori di tipo «sliding window»; oppure possono assumere configurazioni assai complesse se basati su predizione tridimensionale, con predittore commutato (ad esempio in modo adattativo in funzione dell'attività temporale dell'immagine) e codifica a lunghezza variabile.

A parità di efficienza risultano oggi fattibili con minor complessità circuitale codecs basati su algoritmi che abbinano la tecnica delle trasformate, in particolare la Trasformata Coseno Discreta (DCT), a quella della codifica a lunghezza variabile; a parità di complessità, gli studi in atto sembrano confermare che la DCT consente di ottenere rapporti di compressione del bit-rate sostanzialmente più elevati, pur mantenendo invariato l'obiettivo di qualità.

Gli studi, effettuati con lo scopo di valutare le potenzialità delle tecniche basate sulla DCT, hanno portato dapprima allo sviluppo di sistemi di riduzione del bit-rate relativamente semplici, aventi rapporto di compressione pari a circa 5:1; più recentemente, ricorrendo ad algoritmi maggiormente sofisticati, si sono ottenuti risultati soddisfacenti con rapporti di compressione superiori a 10:1.

3. Considerazioni sulla trasformata Coseno Discreta (DCT)

Utilizzando la versione bidimensionale della trasformata, l'immagine viene organizzata in blocchi di dimensioni prefissate, per ciascuno dei quali viene calcolata la trasformata; quest'ultima consiste in blocchi di coefficienti aventi le stesse dimensioni del blocco originario. Circa la trattazione teorica della DCT si rimanda alla bibl. 4.

Sono stati presi in considerazione blocchi di 4x4, 8x8 e 16x16 pixels, tenendo conto che all'aumentare delle dimensioni aumenta l'efficienza dell'algoritmo ma, insieme ad essa, aumentano anche il numero di operazioni e la precisione richiesta per rappresentare i coefficienti trasformati.

Il miglior compromesso consiste nell'impiego di blocchi 8x8 pixels. Il contributo degli 8x8 coefficienti trasformati alla fedeltà con cui l'immagine può venire ricostruita non è costante lungo il singolo blocco; in genere, i termini corrispondenti alla componente continua e alle componenti di bassa frequenza hanno un peso più importante e pertanto la precisione con cui essi vengono rappresentati (espressa in un numero di bit per coefficiente) può essere fatta variare seguendo una procedura adattativa. In base ai risultati finora conseguiti sembra ragionevole una lunghezza di parola aggirantesi sui $10 \div 11$ bit.

La riduzione di ridondanza viene effettuata sfruttando le proprietà statistiche dei coefficienti, la cui distribuzione è caratterizzata da una elevata concentrazione in prossimità del coefficiente che rappresenta la componente continua. In pratica i coefficienti vengono ordinati in modo da convertire la struttura bidimensionale del blocco in un flusso di dati unidimensionale; successivamente, essi vengono codificati secondo un codice di Huffman a lunghezza variabile, scelto in base a considerazioni statistiche svolte su un certo numero di immagini significative.

Si consegue un ulteriore risparmio di informazione trasmessa codificando con una determinata parola eventuali sequenze di coefficienti la cui ampiezza è inferiore ad una certa soglia o trasmettendo una segnalazione di fine blocco qualora tutti i coefficienti a valle di una data posizione siano inferiori alla soglia. Resta sottointeso che in sede di ricostituzione dell'immagine i suddetti coefficienti sono, a tutti gli effetti, trattati come zeri.

In figura 1 è illustrato lo schema di principio di un codificatore operante nel modo intrafield. Poiché la DCT bidimensionale è separabile, si può operare su un blocco 8x8 dapprima lungo le righe, disporre i risultati relativi ad ogni riga ancora lungo una riga e quindi operare sul risultato ottenuto procedendo secondo le colonne; quest'ultima operazione richiede la memorizzazione dei dati su otto righe televisive.

La elaborazione dei coefficienti viene eseguita mediante strutture di calcolo simili alla trasformata discreta veloce di Fourier, con notevole risparmio sul numero di operazioni.

Per aumentare l'efficienza del sistema, in vista di quelle applicazioni in cui è richiesto un elevato rapporto di compressione a parità di qualità globale, occorre sfruttare anche le caratteristiche temporali del segnale; dal punto di vista circuitale ciò comporta l'introduzione di una o più memorie di quadro.

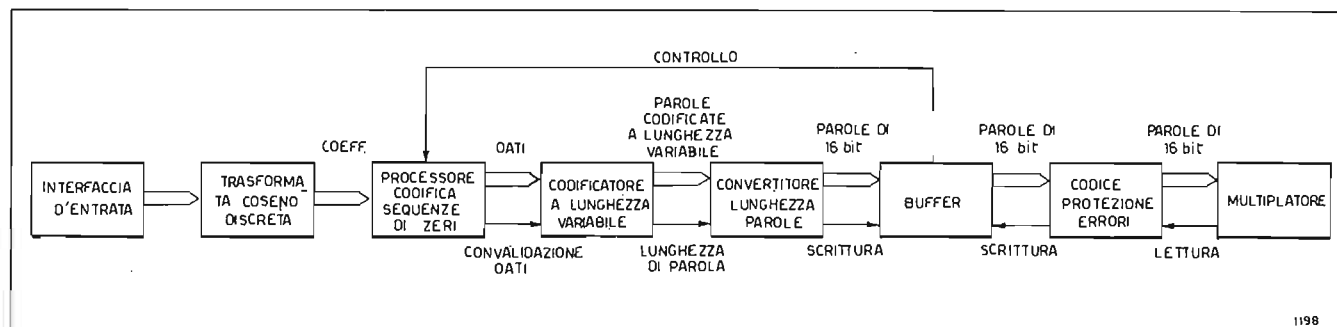



Fig. 1. — Schema di principio del codificatore DCT intrafied.

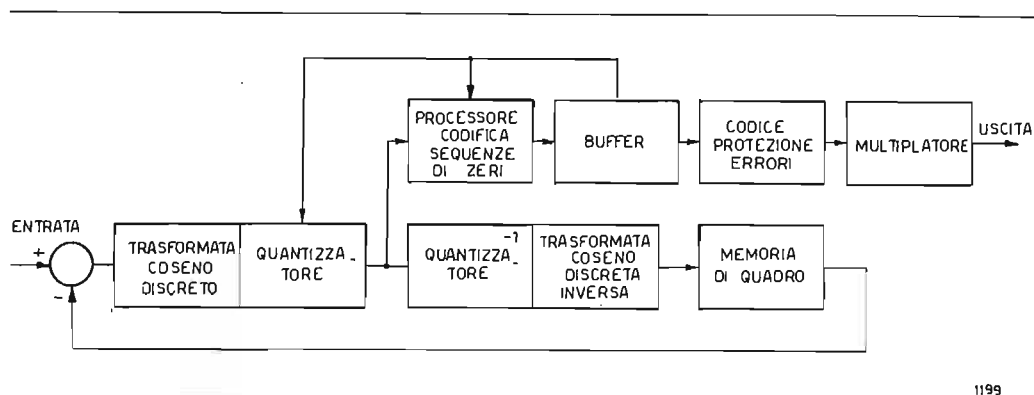



Fig. 2. — Schema di principio del codificatore DCT ibrido.

Una possibile soluzione basata sull'approccio spazio-temporale è illustrato sullo schema di massima di figura 2. La DCT viene ancora calcolata su blocchi di 8x8 termini, rappresentati ciascuno dalle differenze tra elementi di immagine omologhi presi su quadri consecutivi; è immediato valutare i vantaggi di un tale codificatore ibrido nel caso di sequenze di immagini in cui le aree fisse sono presenti statisticamente in misura non trascurabile.

4. Risultati sperimentali

Gli studi sui metodi di riduzione del bit-rate sono stati effettuati operando in simulazione con l'ausilio di un sistema di elaborazione che consente l'acquisizione e la visualizzazione, in tempo reale, di sequenze di immagini televisive a colori aventi durata massima di 90 secondi. Per ragioni pratiche, al fine di contenere i tempi di elaborazione entro limiti ragionevoli si è lavorato essenzialmente su una varietà di sequenze della durata di $3 \div 4$ secondi.

La sperimentazione della soluzione DCT intrafield ha portato ad ottimizzare un modello di codec operante ad un valor medio di bit/campione pari a circa 1.5 [1].

La qualità globale delle immagini è stata ritenuta idonea a soddisfare i requisiti della rete di distribuzione per la quale è accettata sui segnali co-decodificati una degradazione media di 0.5 nella scala CCIR a 5 punti.

Da parte dell'industria nazionale è stato sviluppato un codec basato sul modello suddetto, per la trasmissione a 34 Mbit»s di segnali televisivi numerici a 625 righe (bibl. 5, 6); le relative specifiche nonché i risultati di una campagna di prove soggettive per la valutazione delle prestazioni del sistema sono riportati nella già citata bibliografia 2.

È attualmente oggetto di studio la soluzione DCT ibrida. I risultati finora conseguiti permettono di fare le seguenti considerazioni:

— La riduzione di ridondanza ad un valore medio di 1,5 bit/campione consente di perseguire un obiettivo di qualità che sembra soddisfare i requisiti della rete di contributo; per quest'ultima applicazione si deve prevedere la possibilità di postprocessare il segnale a valle del collegamento, e una degradazione media di 0.5 dopo la connessione in cascata di almeno due codecs.
— L'efficienza dell'algoritmo in corrispondenza di 1,5 bit/campione è ancora sufficientemente elevata; ciò consente di procedere ulteriormente nella riduzione della ridondanza mentre per altri algoritmi, quali ad esempio quelli predittivi, il suddetto traguardo costituisce, allo stato attuale delle conoscenze, un limite inferiore non superabile senza che la qualità venga compromessa in modo inaccettabile.
— La qualità globale ottenuta scendendo a 0.8 bit/campione è stata ritenuta accettabile per applicazioni tipo rete di distribuzione.

[1] Si opera normalmente su segnali televisivi nel formato «in componenti», per i quali il contenuto di informazione alla sorgente è ripartito al 50% tra le componenti di luminanza e differenza di colore. I valori in bit»campione riportati nel presente lavoro si riferiscono a ciascuna delle suddette componenti prese singolarmente.

Gli studi attualmente in corso hanno come obiettivo la riduzione a 0.5 bit/campione; i risultati preliminari sono incoraggianti ma si prevede necessario un intenso lavoro di ottimizzazione degli algoritmi e, soprattutto, occorre procedere ad una verifica approfondita di tali risultati analizzando un numero sufficientemente rappresentativo di sequenze.

5. Riduzione del bit-rate per segnali HDTV

Lo standard di produzione HDTV è tuttora oggetto di acceso dibattito presso i competenti organismi internazionali e pertanto una valutazione del bit-rate del segnale alla sorgente può essere fatta, per il momento, soltanto sulla base di ipotesi. In pratica la discussione è focalizzata sulle proposte illustrate nella tabella I.

Tabella 1

Parametri principali degli standard di produzione HDTV proposti al CCIR.

Numero di righe di scansione (*)	1125 (1035)	1250 (1152)
Frequenza di quadro	30	50
Metodo di scansione	interallac.	progressivo
Campioni per riga attiva (luminanza)	1920	1920
Campioni per riga att. (differenza di colore	960	960

(*) Tra parentesi è indicato il numero di righe attive.

Nel caso indicato nella prima colonna, il numero totale di campioni ammonta a 1920x1035x2=3,97 10^6 campioni/quadro di cui la metà appartengono alla componente di luminanza e l'altra metà alle due componenti differenza di colore, mentre il bit-rate netto (relativo alla sola porzione attiva dell'immagine) vale 0,95 Gbit/s.

Nel caso della seconda colonna si hanno 1920 x 1152 x 2=4.42 10^6 campioni/quadro a cui corrisponde un bit-rate netto di 1,77 Gbit/s.

Poiché alcuni apparati da studio (ad esempio il registratore videomagnetico) si troverebbero al limite delle odierne possibilità tecnologiche dovendo operare con un bit-rate così elevato, viene proposto altresì uno standard intermedio destinato ad essere impiegato nella fase iniziale del servizio; lo standard intermedio differisce da quello della seconda colonna di tabella 1 nel modo di scansione (interallacciata a 25 quadri/s anziché progressiva a 50 quadri/s). Il numero di elementi attivi di immagine per quadro resta invariato mentre, grazie al dimezzamento della frequenza di scansione di quadro, il bit-rate netto si dimezza.

In base ai risultati riportati al punto 4 si possono fare alcune ipotesi circa le possibili prestazioni di una famiglia di codecs per segnali televisivi ad alta definizione, tenendo conto dei requisiti operativi e delle differenti applicazioni previste.

Per soddisfare ai requisiti della rete di contributo sembra necessario attestarsi attorno a valori di compressione aggirantesi sui 1,5 bit/campione; con riferimento agli standard di produzione proposti, si prevedono pertanto i seguenti valori teorici di bit-rate:

- 1,5 x 3,97 x 30=178,8 Mbit/s (sistema 1125/30)
- 1,5 x 4,42 x 25=165,8 Mbit/s (sistema 1250/50 interallacciato)
- 1,5 x 4,42 x 50=331 Mbit/s (sistema 1250/50 progressivo)

Sono in corso studi per valutare se, mediante un ulteriore affinamento degli algoritmi, è possibile ricondurre i primi due sistemi a funzionare entro il livello gerarchico dei 140 Mbit/s.

Per applicazioni concernenti la distribuzione (attraverso reti terrestri o via satellite) occorre tener conto che i parametri dello standard di produzione sono dimensionati in modo da assicurare quei margini operativi che solitamente risultano necessari per consentire la postprocessabilità del segnale; i valori di bit-rate precedentemente citati appaiono pertanto essere alquanto severi.

In particolare, la scansione progressiva rivela la sua utilità in quelle operazioni di trattamento dell'immagine ove è richiesta una migliore rappresentazione delle sue caratteristiche temporali; tuttavia, a causa della capacità trasmissiva che si verrebbe ad impegnare, non si può considerare praticabile un tale approccio tecnico nei servizi di distribuzione.

Gli obiettivi di qualità per la diffusione via satellite di segnali HDTV in canali a banda stretta (24÷27 MHz) sono oggetto di discussione al CCIR (bibl. 7). Con riferimento alle considerazioni svolte al punto 4, si può ritenere che un sistema di riduzione del bit-rate il quale realizzi una compressione a 0.8 bit/campione possa essere ritenuto conforme alle esigenze del servizio di diffusione.

Con riferimento al suddetto livello di compressione si presume che un codec sviluppato per applicazioni alla rete di distribuzione possa funzionare nella seguente classe di bit-rate netti:

- 0,8 x 3,97 x 30=95,3 Mbit/s (sistema 1125/30, interallacciato)
- 0,8 x 4,42 x 25=88 Mbit/s (sistema 1250/25 interallacciato).

I suddetti valori risultano, tuttavia, piuttosto elevati per i canali di trasmissione che, realisticamente, si possono prevedere attuabili nel prossimo futuro. Nel caso della trasmissione via satellite occorre fare i conti, in particolare, con il rumore di canale che, generalmente, richiede l'impiego di adeguati codici di protezione con inevitabile aumento del bit-rate totale.

Nel caso in cui gli studi in corso confermino la fattibilità, seppure in prospettiva futura, di un sistema di compressione a 0,5 bit/campione si otterrebbero i seguenti valori netti:

- 0,5 x 3,97 x 30=59,5 Mbit/s (sistema 1125/30, interallacciato)
- 0,5 x 4,42 x 25=55 Mbit/s (sistema 1250/25, interallacciato).

In entrambi i casi verrebbe lasciato sufficiente margine per sviluppare un multiplex operante nella fascia dei 60 ÷ 70 Mbit/s, con adeguata capacità trasmissiva per i codici di protezione e per tutte le restanti informazioni che si riterrà necessario inserire (audio, dati ed eventuali segnali ausiliari).

6. Fattibilità hardware

Sulla base della complessità circuitale risultante dalla realizzazione del codec a 34 Mbit/s per segnali a 625 righe, è possibile fare le seguenti considerazioni circa lo sviluppo di un analogo apparato per segnali HDTV, operante secondo gli stessi principi:

— I dati provenienti dall'interfaccia d'entrata vengono scritti in una memoria tampone veloce, organizzati in modo opportuno e distribuiti ad un certo numero (4 ÷ 5) di codificatori DCT operanti in parallelo. L'unica funzione comune rimane il processore controllato dal riempimento del «buffer» di trasmissione.
— A parte gli stadi iniziale e finale, tutte le altre funzioni del codec (DCT, quantizzazione, codifica a lunghezza variabile, protezione dagli errori e memoria di quadro, nel caso della versione ibrida) operano a velocità ridotta e possono pertanto venire svolte da componentistica HCMOS.
— L'impiego delle tecnologie HCMOS consente un elevato grado di integrazione (particolarmente importante se si tiene conto che il numero globale di gates per un codec in versione HDTV può essere stimato 4 ÷ 5 volte superiore a quello di un corrispondente codec per segnali convenzionali).
— Si può ritenere che la dissipazione globale di un codificatore HDTV venga contenuta nella misura di 70 ÷ 100 watt contro gli attuali 15 watt del codificatore a 34 Mbit/s.

7. Conclusioni

La codifica numerica dei segnali televisivi ad alta definizione è caratterizzata da un bit-rate, alla sorgente, aggirantesi sul Gbit/s ed oltre. Il problema di ricondurre tale velocità a valori compatibili con la capacità trasmissiva dei canali realizzabili oggi, o nell'immediato futuro, è oggetto di intensi studi.

Le attuali tecnologie consentono di sviluppare algoritmi ad elevata efficienza, mediante i quali la quantità di informazione utile da trasmettere viene sostanzialmente ridotta sfruttando le ridondanze intrinseche nel segnale di sorgente.

La sperimentazione, effettuata principalmente in Italia, sulle potenzialità della Trasformata Coseno Discreta ha permesso finora di verificare la fattibilità di sistemi in cui l'informazione viene compressa di un rapporto variabile tra circa 5:1 e circa 10:1, pur restando i livelli qualitativi conformi alle specifiche imposte dalle esigenze del servizio.

Nelle reti di contributo, ove il segnale deve sottostare a particolari requisiti dettati dalla normale prassi produttiva, è realistico prevedere adeguato un bit- rate netto aggirantesi sui 160 ÷ 180 Mbit/s; tenendo conto delle possibili gerarchie trasmissive risulterebbe necessario ricorrere ad un multiplex da 2x140 Mbit/s. Si tenga tuttavia presente che lo sviluppo di una rete operante a tale velocità non sembra realizzabile se non a tempi lunghi; un obiettivo da perseguire con urgenza appare pertanto quello di migliorare l'efficienza degli attuali algoritmi in modo da consentire lo sviluppo di apparati il cui funzionamento rientri nel livello gerarchico dei 140 Mbit/s.

Circa la distribuzione, in particolare la diffusione via satellite, i risultati della sperimentazione finora condotta hanno dimostrato che un bit-rate aggirantesi sui 90 ÷ 100 Mbit/s netti, da assegnare all'immagine codificata, è sufficiente a garantire gli obiettivi di qualità proposti.

Sembrano tuttavia esistere sufficienti margini per portare, con un maggior affinamento degli algoritmi, e in una prospettiva a più lungo termine, la compressione dell'informazione a livelli più spinti; in particolare sembra possibile ottenere velocità di codifica di poco inferiore a 60 Mbit/s netti.

Si ritiene opportuno sottolineare il carattere del tutto preliminare dei risultati; sono in corso studi per verificare l'attendibilità di tale ipotesi che, se si dimostrasse fattibile, consentirebbe di realizzare un sistema di trasmissione di televisione ad alta definizione ad un bit-rate operante nella fascia 60 ÷ 70 Mbit/s, comprensivi della codifica d'errore, dell'audio e dei segnali ausiliari.

(3711)

BIBLIOGRAFIA

1. - CCIR - Doc. 11/88 (1986-1990): *Transmission of television signals on digital networks at H2 CCITT access level.*

2. - CCIR - Doc. IWP 11/7 (1986-1990): *DCT coding of 4:2:2 television signals for transmission at 34 Mbit/s: subjective tests.*

3. - CCIR - Report 1089: *Bit-rate reduction for digital television signals.*

4. - MAKHOUL J.: *A fast cosine transform in one and two dimensions.* «IEEE Trans. On Acoustics, Speech, Signal processing», Vol., ASSP-28, n. 1, Febbraio 1980.

5. - CESA, MODENA, SICURANZA: *2D DCT intrafield coding on high-quality TV signals.* Conference on DSP, Firenze 1987.

6. - CUCCHI, MODENA: *Component coding of TV signal using intrafield DCT transform: simulation results and considerations about hardware complexity.* 5° World communication Forum, Geneve, ottobre 1987.

7. - CCIR - Doc. 10-11S/91 (1986-1990): *High definition Television by Satellite.* CCIR Report 1075.

CODIFICA DEL SEGNALE TELEVISIVO NUMERICO:
METODI DI RIDUZIONE DELLA RIDONDANZA

ROBERTO DEL PERO*

SOMMARIO — Nel presente lavoro si analizza qual'è l'informazione in sovrappiù presente nel segnale televisivo (ridondanza) eliminando la quale non si ha peggioramento oggettivo o soggettivo nella qualità dell'immagine. Si fornisce, in seguito, una visione generale dei metodi attualmente utilizzati per ridurre l'informazione ridondante con particolare attenzione ai metodi basati sulla codifica predittiva e a quelli utilizzanti le trasformazioni ortogonali. Tra queste ultime, viene analizzata la Trasformata Coseno Discreta (DCT) bidimensionale, che permette di ottenere rapporti di compressione nel flusso binario (bit rate) variabili tra circa 1:6 e 1:12 senza perdita apprezzabile di qualità. Analisi sull'efficienza di schemi di codifica basati sulla DCT sono in fase di esecuzione presso il Centro Ricerche RAI di Torino tramite simulazioni su calcolatore; sulla base dei risultati di detti studi è in avanzata fase di sviluppo un codec operante a 34 Mbit/s per segnali TV standard.

SUMMARY — ***Digital TV signal coding: methods of redundancy reduction.*** *This work analyses the nature of the redundant information present in the television signal, whose elimination causes neither objective nor subjective worsening of the picture quality. A general view of methods at present utilized to reduce the redundant information is then given, with particulary attention to methods based on the predictive coding and methods which utilize the orthogonal transforms. Among them the two dimensional discrete cosine transform (DCT) is analyzed, by which compression ratios in the bit rate between 1:6 and 1:12 is obtained without perceptible loss of quality. Analyses on the efficiency of advanced coding scheme based on the DCT are evaluated at the RAI Research Centre by computer simulations; meanwhile on the results of these studies is based the design and implementation of a codec operating at 34 Mbit/s for standard TV signals.*

1. Introduzione

La codifica del segnale televisivo numerico secondo la Raccomandazione CCIR 601 richiede un flusso binario (bit rate) di 216 Mbit/s: 108 Mbit/s per il segnale di luminanza campionato a 13,5 MHz con 8 bit per campione e 54 Mbit/s per ciascuna delle due componenti di crominanza campionate a 6,75 MHz e codificate con 8 bit per campione.

Una prima riduzione di questo bit rate può essere ottenuta trasmettendo la sola porzione attiva dell'immagine, raggiungendo, così, un valore minimo pari a 165,9 Mbit/s. Tale riduzione è tuttavia insufficiente per applicazioni nelle quali l'ottimizzazione del servizio è legata alla larghezza di banda del segnale da trasmettere oppure nelle quali esistano problemi dovuti a precedenti pianificazioni di banda dei canali di trasmissione.

Questo problema è ancora più sentito in applicazioni riguardanti la televisione ad alta definizione (HDTV) per la quale i bit rate in gioco sono superiori a 1 Gbit/s.

Riduzioni di bit rate sicuramente superiori a quelle sopra citate sono ottenibili per mezzo della riduzione della ridondanza presente in una qualsiasi immagine televisiva. Col termine ridondanza si intende quella parte di informazione non strettamente necessaria alla ricostruzione dell'informazione trasmessa.

2. Riduzione della ridondanza

Nel caso di sorgente televisiva si possono distinguere tre tipi di informazione eliminabile: l'informazione ridondante statisticamente, l'informazione irrilevante e l'informazione ridondante spettralmente.

L'informazione ridondante statisticamente è quella che può essere eliminata in trasmissione senza alcuna perdita di qualità nell'immagine ricostruita, in quanto tale informazione può essere recuperata completamente tenendo conto delle mutue relazioni esistenti tra elementi di immagine (pixel) adiacenti (correlazione spazio-temporale tra pixel adiacenti).

L'informazione irrilevante è legata alla ridondanza psicovisiva: parte dell'informazione è eliminabile in quanto gli effetti visivi generati dalla sua mancanza sono irrilevanti a causa del modo di lavorare del sistema visivo umano (ad esempio la sensibilità dell'occhio diminuisce al crescere della frequenza spaziale). L'eliminazione di que-

(*) Ing. Roberto Del Pero del Centro Ricerche RAI (Torino).
Dattiloscritto pervenuto alla Redazione il 26 gennaio 1989.

sta parte dell'informazione crea un degradamento non visibile o scarsamente visibile nell'immagine ricostruita.

Infine, la ridondanza spettrale deriva dalla struttura del campionamento dell'immagine: utilizzando opportune strutture di campionamento adattate alle effettive caratteristiche spettrali del segnale televisivo, è possibile operare con frequenze di campionamento inferiori a quelle adottate attualmente senza perdita di qualità dovuta a fenomeni di interferenza (aliasing).

2.1 RIDUZIONE DELLA RIDONDANZA STATISTICA

La riduzione della ridondanza statistica è una operazione del tutto reversibile che richiede, per poter essere effettuata al meglio, una conoscenza della sorgente in termini di quantità di informazione generata. La grandezza che descrive questa quantità è l'entropia della sorgente definita come informazione media generata dalla sorgente. Si può dimostrare che essa rappresenta il minimo numero di bit necessari per rappresentare completamente l'uscita della sorgente senza distorsioni. Tale minimo è raggiungibile utilizzando una codifica a lunghezza variabile (VLC) tramite la quale si associano ai messaggi meno probabili parole più lunghe: naturalmente è fondamentale una conoscenza corretta della distribuzione statistica dei messaggi (nel nostro caso un messaggio è il valore della luminanza o della crominanza associata al pixel).

Metodo ottimo di codifica a lunghezza variabile è quello di Huffman: l'algoritmo utilizzato richiede la conoscenza della probabilità associata ai diversi messaggi e dei legami statistici esistenti tra essi, ed esso sarà tanto più efficiente (eliminerà tanto meglio la ridondanza statistica) quanto più i valori di probabilità adottati saranno simili a quelli reali; questo significa che, mentre per ottimizzare il codice di Huffman associato ad una sorgente che genera messaggi statisticamente indipendenti è sufficiente conoscere la statistica del primo ordine associata a quella sorgente, per messaggi statisticamente dipendenti, quali quelli caratteristici di una sorgente televisiva, è necessario lavorare con statistiche di ordine superiore che tengano conto delle relazioni esistenti tra i messaggi generati.

Nel caso di immagini televisive i messaggi non sono assolutamente indipendenti dal punto di vista statistico in quanto ciascun pixel è correlato, più o meno strettamente, con tutti i pixel ad esso adiacenti sia spazialmente che temporalmente: per ridurre al meglio la ridondanza occorre quindi una conoscenza statistica di ordine molto elevato, cosa non attuabile dal punto di vista del calcolo.

Si può superare questo ostacolo sottoponendo l'immagine ad una trasformazione che riduca la correlazione tra i vari elementi dell'immagine e che sia, nello stesso tempo, reversibile, permetta cioè una ricostruzione perfetta dell'originale: la conoscenza della statistica del primo ordine dell'immagine così ottenuta permette di calcolare un codice di Huffman che, applicato all'immagine trasformata, si adatta meglio alla sorgente effettiva.

Attualmente le trasformazioni adottate per ridurre la correlazione tra gli elementi di una immagine sono essenzialmente di due tipi: la *codifica predittiva* e le *trasformazioni ortogonali*.

Un'altra caratteristica delle immagini numeriche che può essere sfruttata per la riduzione della ridondanza statistica è quella per cui pixel di valore uguale hanno una probabilità relativamente alta di essere adiacenti; perciò è conveniente codificare non il singolo pixel, ma la sequenza di pixel uguali con un'unica parola di codice (*codifica di sequenze o run length encoding*).

2.2 RIDUZIONE DELLA IRRILEVANZA

La riduzione della irrilevanza è una operazione non reversibile in quanto tende ad eliminare una parte di informazione non ridondante statisticamente; essa introduce una distorsione oggettiva nell'immagine trattata, in modo tale, però, da rendere minimo l'effetto soggettivo che tale distorsione ha per l'osservatore.

Essenzialmente sono due le possibili operazioni adottate per ottenere questo: la *quantizzazione* e la *codifica psicovisiva*.

La *quantizzazione* opera una riduzione del numero di bit utilizzati per rappresentare ciascun campione: con questa operazione viene introdotto del rumore sull'immagine. Scegliendo opportunamente l'ampiezza degli intervalli di quantizzazione e gli associati valori di restituzione, è possibile minimizzare l'effetto soggettivo di questo rumore, ad esempio inserendolo soltanto in zone di immagine ad alto livello di luminanza oppure associandolo in misura maggiore alle crominanze rispetto alla luminanza (l'occhio è più sensibile a difetti su quest'ultima componente).

La *codifica psicovisiva* si basa sull'osservazione che la sensibilità dell'occhio umano è funzione della frequenza spaziale del segnale stimolante: essa decresce al crescere della frequenza. Se è possibile introdurre delle distorsioni sull'immagine in modo che ne siano coinvolte soltanto le componenti di alta frequenza spaziale, a parità di distorsione oggettiva, si riesce ad ottenere una minore distorsione soggettiva.

2.3 RIDUZIONE DELLA RIDONDANZA SPETTRALE

Come la riduzione dell'irrilevanza, anche la riduzione della ridondanza spettrale è una operazione irreversibile. Viene normalmente effettuata tramite un prefiltraggio multidimensionale (spaziale o spazio-temporale), che limita la banda dell'immagine d'ingresso in modo da evitare fenomeni di aliasing, seguito da un sottocampionamento (solo un sottoinsieme di campioni viene preso in considerazione). Scegliendo opportunamente la struttura del campionamento ed il tipo di prefiltraggio, è possibile minimizzare la distorsione introdotta.

Queste operazioni possono essere effettuate sull'intera immagine da trasmettere oppure soltanto su aree particolari di essa nelle quali i difetti siano meno percettibili da parte dell'osservatore: normalmente, tali zone sono quelle che presentano movimento al di sopra di una certa soglia. L'approccio zonale richiede una accurata caratterizzazione dell'immagine dal punto di vista della segmentazione in aree con e senza i requisiti richiesti per l'applicazione dell'algoritmo, cosa che appesantisce notevolmente l'algoritmo dal punto di vista dei calcoli da eseguire.

Il difetto che normalmente insorge nelle immagini così trattate è una perdita più o meno importante di definizione dovuta al prefiltraggio anti-aliasing, difetto appunto meno visibile nelle zone dell'immagine in movimento.

Strutture di campionamento attualmente prese in considerazione sono il *line-quincunx* ed il *field-quincunx* che operano rispettivamente nel campo spaziale ed in quello spazio-temporale. Per quanto riguarda il prefiltraggio, vengono utilizzati filtri numerici multidimensionali sia spaziali che spazio-temporali.

3. Codifica predittiva

Gli algoritmi di codifica predittiva si basano sull'osservazione che, data l'alta correlazione esistente tra pixel adiacenti, è possibile prevedere con sufficiente precisione il valore di un pixel a partire da un insieme di pixel adiacenti già noti. Sfruttando questo fatto, gli algoritmi di codifica predittiva non operano sul valore effettivo del pixel, bensì sulla differenza tra questo valore ed il valore previsto (errore di predizione): il vantaggio risultante è che l'errore di predizione relativo ai vari pixel è maggiormente scorrelato rispetto al valore assoluto dei pixel stessi. Quest'ultima affermazione è valida se l'algoritmo di predizione adottato è tale da minimizzare l'errore di predizione.

In generale la predizione del valore assunto da un pixel si ricava utilizzando una funzione che coinvolge i valori di un insieme di pixel adiacenti (spazialmente e temporalmente) a quello in esame: il problema da risolvere in questo tipo di codifica è quello della determinazione della funzione di predizione ottima che minimizza l'errore di predizione. Purtroppo non esiste in realtà una funzione predittiva ottima in assoluto, ma questa dipende fortemente dalle caratteristiche dell'immagine da trattare, dai dettagli presenti, dal movimento, dai contrasti di luminanza, dai contorni e così via; questa limitazione è superabile in parte adottando dei predittori adattativi nei quali la legge di stima del valore del pixel viene adattata alla zona di immagine in esame in modo da ottimizzare localmente la predizione.

In generale si possono distinguere due tipi di predittori, quelli operanti secondo *leggi inter* e quelli che si basano su *leggi intra*. Le funzioni che operano in modo intra sono quelle che utilizzano come dati di base i valori di pixel appartenenti allo stesso semiquadro (field) in esame, mentre le leggi inter lavorano su pixel appartenenti a semiquadri (inter-field) o a quadri (frame) (inter-frame) precedenti.

Le funzioni intra sono chiaramente più semplici dal punto di vista della realizzazione, ma offrono risultati mediocri nel caso in cui si abbia a che fare con sequenze di immagini con poco movimento: infatti, in sequenze di questo tipo, è molto elevata la correlazione tra pixel spazialmente coposizionati ma appartenenti a field/frame diversi, correlazione che non viene eliminata agendo solo su un field alla volta.

Viceversa i modi inter si dimostrano inadeguati nel trattare sequenze di immagini ad alto movimento: per movimenti rapidi, infatti, esiste una bassa correlazione tra il pixel in esame ed i pixel ad esso adiacenti ma su field/frame precedenti mentre, a causa dell'operazione di integrazione effettuata dal sistema di ripresa, aumenta la correlazione esistente tra pixel adiacenti appartenenti allo stesso field e posti nella direzione del movimento.

Come già detto, un modo per limitare problemi di questo tipo è quello di rendere adattative le leggi di predizione segmentando l'immagine in zone ad alto movimento da trattare in modo intra-field, in zone statiche da elaborare in modo inter-frame ed in zone quasi-statiche alle quali applicare il modo inter-field.

Ulteriore miglioramento si può avere con l'introduzione della tecnica della *compensazione del movimento*. Questa tecnica tiene conto del movimento presente nell'immagine ricavando uno o più vettori di spostamento, i quali vengono utilizzati per determinare quali pixel, nei field precedenti, sono i più correlati con i pixel della zona in esame e quindi quali sono i pixel da utilizzare per effettuare la predizione in modo inter-field o inter-frame.

Una volta ottenuti i valori di errore di predizione tramite i metodi analizzati prima, questi vengono quantizzati opportunamente (in modo non uniforme) e quindi codificati. Le caratterestiche del segnale di errore di predizione sono tali per cui, effettuando una codifica normale (non entropica, cioè non utilizzando codici a lunghezza variabile), si ha una riduzione di circa 2 bit/pixel per ogni componente (pari a circa 60 Mbit/s complessivi) nel bit rate necessario alla trasmissione, mentre ulteriori miglioramenti si hanno sfruttando la bassa correlazione tra gli errori di predizione codificandoli con una codifica a lunghezza variabile (VLC) opportuna. In quest'ultimo caso è necessario prevedere in cascata al codificatore un buffer che faccia da tampone tra il flusso di dati a velocità variabile in uscita dal sistema e il flusso costante normalmente richiesto sul canale di trasmissione; per evitare situazioni critiche quali il riempimento o lo svuotamento del buffer in seguito alla codifica di immagini particolari, è necessario inoltre controllarne l'occupazione e, tramite retroazione, agire opportunamente sul blocco di quantizzazione e codifica.

Aspetto negativo dei sistemi a codifica predittiva è quello della complessità realizzativa necessaria per ottenere riduzioni significative di bit rate: è indispensabile infatti lavorare con predizione adattativa e compensazione del movimento.

4. Trasformazioni ortogonali

Nelle trasformazioni ortogonali l'immagine viene sottoposta ad una trasformazione lineare ed invertibile tramite la quale i singoli pixel vengono sostituiti da coefficienti la cui mutua correlazione risulta minore. Normalmente, il calcolo dei coefficienti trasformati viene effettuato non sull'intera immagine contemporaneamente, ma su blocchetti di minore dimensione in quanto risulta più agevole la realizzazione dell'algoritmo di trasformazione; la contropartita è una riduzione non ottimale della correlazione in quanto si considerano solo i legami esistenti tra i pixel all'interno del blocchetto senza tenere in conto la correlazione esistente tra questi ultimi e quelli all'esterno.

Dal punto di vista matematico l'operazione di trasfor-

mazione è una proiezione di funzioni continue o discrete lungo funzioni mutuamente ortonormali.

Tra le varie leggi che possono regolare la trasformazione pixel-coefficiente, quella ottima è la Karhunen-Loève (KLT): essa fornisce in uscita coefficienti del tutto scorrelati tra loro, ma ha il difetto di dipendere dall'immagine alla quale viene applicata, in quanto richiede, per la determinazione delle funzioni di base, la conoscenza della statistica dell'immagine da trasformare. L'utilizzo di una statistica media, calcolata ad esempio su un certo numero di immagini considerate standard, semplifica l'algoritmo, ma fornisce risultati subottimi. La trasformata KLT non viene quindi utilizzata praticamente, ma è utile come riferimento nella valutazione delle prestazioni delle altre leggi di trasformazione.

Altre leggi di trasformazione che possono avere un certo interesse sono la trasformata di Fourier discreta (DFT), la trasformata di Walsh-Hadamard (WHT) e la trasformata coseno discreta (DCT). La prima ha prestazioni simili alla KLT soltanto nel caso in cui si operi sull'intera immagine contemporaneamente, e presenta, quindi, un carico di calcolo non ammissibile; la WHT è viceversa di facile sviluppo ma fornisce prestazioni inferiori rispetto alle altre leggi di trasformazione. La DCT è attualmente oggetto delle analisi più attente in quanto fornisce prestazioni molto vicine alla KLT pur non presentando eccessive complicazioni dal punto di vista della realizzazione.

5. Trasformata Coseno Discreta

La Trasformata Coseno Discreta (DCT) è la trasformazione ortogonale attualmente più utilizzata nel campo della riduzione del bit rate. Confrontata con altre trasformazioni ortogonali, essa è quella che più si avvicina alla trasformata ottima di Karhunen-Loève per un'ampia classe di segnali.

5.1. Formulazione matematica della DCT

Dal punto di vista dei calcoli, è dimostrato che gli M coefficienti DCT relativi a M campioni del segnale si possono ottenere utilizzando la trasformata di Fourier discreta (DFT) estesa ai $2M$ punti ottenuti effettuando una estensione pari del segnale di partenza; è così possibile utilizzare algoritmi di tipo veloce (fast) derivati da quelli messi a punto per il calcolo della DFT che permettono di ridurre notevolmente il numero di operazioni necessarie per il calcolo dei coefficienti. Con uno stesso numero di operazioni è possibile ottenere anche la trasformazione inversa.

La trasformata coseno discreta di una sequenza di campioni $X(m)$ con $m = 1, 2, \dots (M-1)$, è definita come:

$$G_x(k) = 2 \cdot \sum_{m=0}^{M-1} X(m) \cos \frac{(2m+1)k\pi}{2M}$$

con $k = 0, 1, \dots, (M-1)$

dove $G_x(k)$ è il k-esimo coefficiente della DCT.

Per quanto riguarda il legame esistente tra la DCT e la DFT è possibile dimostrare che, chiamato $Y(k)$ il k-esimo coefficiente della DFT estesa a $2M$ punti e $C(k)$ quello della DCT, vale la seguente relazione:

$$Y(k) = W_{2M}^{-\frac{k}{2}} C(k) \quad \text{con} \quad W_M^P = e^{-j2\pi P/M}$$

La trasformazione inversa (IDCT) viene definita come:

$$X^{\wedge}(m) = \frac{1}{2M} G_x(0) \cdot \frac{1}{M} \sum_{k=1}^{M-1} G_x(k) \cos \frac{(2m+1)k\pi}{2M}$$

con

$$m = 0, 1, \dots, (M-1)$$

dove $X^{\wedge}(m)$ è la sequenza ricostruita.

Le definizioni viste sopra si riferiscono al caso monodimensionale. Prestazioni migliori dal punto di vista della riduzione della correlazione, si ottengono operando su trasformazioni pluri-dimensionali (normalmente bi- o tridimensionali), in quanto esse tendono ad eliminare la correlazione esistente tra pixel adiacenti in più direzioni.

Nel caso bidimensionale, la trasformata coseno può essere vista come operazione tra matrici: chiamando $[f(x, y)]$ la matrice contenente i campioni relativi alla zona da trasformare (di dimensione $N \times N$), $[F(u, v)]$ quella dei coefficienti ottenuti tramite la trasformazione, C la matrice di trasformazione e C^T la sua trasposta, si ha la relazione:

$$[F(u, v)] = C[f(x, y)]\, C^T$$

e, per l'operazione di antitrasformazione:

$$[f^{\wedge}(x, y)] = C^T[F(u, v)]\, C$$

con $[f^{\wedge}(x, y)]$ matrice contenente i campioni ricostruiti.

La matrice di trasformazione C è definita come:

$$C = N^{-1/2}[C_{rc}]$$

con

$$C_{rc} = \begin{cases} 1 & \text{per } r = 0 \\ \sqrt{2} \cdot \cos \dfrac{(2c+1)r\pi}{2N} & \text{per } r = 1, \dots, N-1 \end{cases}$$

con

$$c = 0, 1, \dots, N-1$$

(definizione canonica con coefficienti moltiplicativi suddivisi in modo uguale nella trasformata e nella antitrasformata).

5.2 Significato spettrale dei coefficienti

Visti i legami esistenti tra la DCT e la DFT, è possibile dare un significato spettrale ai coefficienti ottenuti tramite la trasformazione in esame.

Analizzando il caso bidimensionale, ogni elemento della matrice $[F(u, v)]$ è circa proporzionale all'ampiezza di una specifica armonica di frequenza spaziale, armonica crescente al crescere degli indici matriciali; in particolare, il coefficiente di indici (0,0) è circa proporzionale alla componente continua presente nella zona di immagine sotto-

posta alla trasformazione, i coefficienti di indici (0, v), con v diverso da 0, sono circa proporzionali alle ampiezze delle armoniche con frequenze puramente orizzontali, mentre ai coefficienti di indici (u, 0), con un u diverso da 0, sono associate le frequenze puramente verticali. I restanti coefficienti di indici (u, v), con u e v diversi da 0, sono legati alle ampiezze delle armoniche con frequenze puramente diagonali.

In appendice sono riportati alcuni esempi di trasformazione applicata a blocchetti di dimensione 8 × 8 nei quali sono evidenziati i legami tra la struttura spaziale e la posizione dei coefficienti DCT generati.

6. Utilizzo della DCT

Come detto in precedenza, il peso dal punto di vista della realizzazione e dei tempi di calcolo cresce al crescere delle dimensioni dei blocchetti di pixel sui quali applicare la trasformazione; d'altro canto, le prestazioni delle trasformate, quanto a riduzione della correlazione, migliorano al crescere del numero di pixel presi in esame. Per quanto riguarda la DCT, si è trovato che il miglior compromesso tra le due tendenze si ha considerando blocchetti di dimensione 8 × 8 pixel.

L'operazione di trasformazione applicata ai suddetti blocchetti dà origine ad un numero di coefficienti pari al numero di pixel coinvolti nell'operazione, coefficienti che, però, presentano tra loro una minore correlazione. Applicando, quindi, ad essi una codifica a lunghezza variabile (ad esempio un codice di Huffman) si ottiene una maggiore riduzione di ridondanza statistica rispetto a quella che si avrebbe operando sui pixel reali.

L'andamento dei valori dei coefficienti risulta, mediamente, fortemente non uniforme nel senso che, normalmente, solo pochi coefficienti risultano avere valori relativamente elevati mentre sono numerosi quelli a valore basso o addirittura nullo. Questa non uniformità è basilare per ottenere una buona riduzione di bit rate tramite codifica a lunghezza variabile ed inoltre permette di utilizzare con successo algoritmi di tipo codifica delle sequenze applicati alle sequenze di zeri; infatti, scegliendo opportunamente l'ordine con il quale trasmettere i coefficienti di un blocchetto è possibile avere lunghe sequenze di valori nulli, sequenze codificabili con un'unica parola di codice.

Analizzando i coefficienti a valore basso, si può notare che essi hanno indici posizionali alti e quindi sono relativi ad armoniche di alta frequenza spaziale, soprattutto diagonali, alle quali l'occhio umano è relativamente poco sensibile. Si può, quindi, pensare di azzerare questi coefficienti senza per questo peggiorare l'aspetto soggettivo dell'immagine, ma rendendo ancora meno uniforme la distribuzione e soprattutto aumentando il numero e la lunghezza delle sequenze di zeri. Analogamente è possibile introdurre una quantizzazione più grossolana sui coefficienti di alta frequenza e su tutti i coefficienti relativi alle componenti di crominanza.

Anche in questo caso l'utilizzo di codici a lunghezza variabile rende necessaria la presenza di un buffer di uscita per rendere costante il flusso di dati sul canale di trasmissione.

7. Miglioramenti

Le prestazioni ottenibili per mezzo della DCT possono essere ulteriormente migliorate considerando anche la correlazione esistente tra pixel appartenenti a semiquadri diversi.

Un primo tipo di approccio consiste nell'utilizzo della DCT tridimensionale: non si lavora più su blocchetti di pixel appartenenti ad un unico semiquadro, ma si considerano anche i pixel corrispondenti posti sui semiquadri precedenti. In questo modo i coefficienti ottenuti presentano una minore correlazione non soltanto spaziale ma anche temporale ed aumenta il numero e la lunghezza delle sequenze di zeri.

La DCT tridimensionale può essere scomposta in due operazioni distinte: una trasformata bidimensionale all'interno dei semiquadri seguita da una trasformata monodimensionale nella direzione temporale, la prima, normalmente, eseguita su blocchi di 8 × 8 pixel, la seconda su un numero limitato di semiquadri (normalmente 4) per diminuire la memoria necessaria.

Una soluzione attualmente più praticabile è quella definita *ibrida* nella quale si continua ad utilizzare la DCT bidimensionale applicata però di volta in volta a blocchetti di pixel (modo di lavoro intrafield) oppure a blocchetti contenenti gli errori di predizione relativi (modi di lavoro interfield o interframe).

In sistemi di questo tipo la funzione di predizione utilizzata è molto semplice: infatti, come valore stimato si utilizza il valore del pixel coposizionato sul semiquadro o sul quadro precedente.

La scelta tra modi inter e intra viene effettuata analizzando il valore di una funzione di attività legata in modo opportuno alla quantità di dettaglio e di movimento presente nel blocchetto in esame in modo tale che blocchetti con poco movimento vengono codificati con i modi inter, mentre zone con alta attività lavorano in modo intra.

È possibile anche in questi sistemi introdurre algoritmi di compensazione del movimento come già visto per la codifica predittiva.

8. Conclusioni

Gli algoritmi basati sulla DCT sembrano rappresentare, attualmente, il miglior compromesso tra efficienza e complessità della struttura dell'hardware (un primo prototipo di trasmettitore basato sulla DCT bidimensionale non ibrida ha richiesto per la sua costruzione 5 schede di formato 16 × 23 cm).

Da simulazioni su calcolatore eseguite presso il Centro Ricerche RAI, l'algoritmo basato sulla DCT ibrida si è dimostrato molto promettente, permettendo di ottenere fattori di compressione del bit rate compresi tra 1:6 e 1:12 senza deterioramento apprezzabile dell'immagine; alla luce di tali risultati è stato avviato, in collaborazione con una importante azienda nazionale nel campo delle telecomunicazioni, lo sviluppo hardware di co-decodificatori operanti a bit rate dell'ordine di 30 Mbit/s utilizzabili per applicazioni di tipo distribuzione con immagini in formato standard 625 righe per quadro (nei prossimi mesi saranno disponibili i primi prototipi).

In vista di utilizzazioni a più basso bit rate oppure nel campo della televisione ad alta definizione, sono in fase di sviluppo programmi per la simulazione di nuovi sistemi, sempre basati sulla DCT ibrida, nei quali vengono implementati algoritmi più sofisticati per la riduzione della ridondanza psicovisiva: grazie a questi accorgimenti, si prevede il raggiungimento di bit rate dell'ordine di 140 Mbit/s in applicazioni HDTV mantenendo una qualità forse adatta anche alle reti di contributo (si definisce rete di contributo una rete di collegamento a livello nazionale o internazionale ai capi della quale deve essere possibile effettuare senza perdita di qualità una qualsiasi elaborazione delle immagini ricevute).

APPENDICE

In tabella I sono riportate le matrici $[F(u, v)]$ dei coefficienti relativi alla trasformata coseno discreta bidimensionale applicata a dieci strutture spaziali, di formato 8×8, particolarmente utili per comprendere il comportamento della DCT.

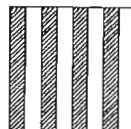
I valori numerici presentano delle differenze rispetto a quelli calcolabili mediante le formule date nel paragrafo 5.1 in quanto l'algoritmo utilizzato non prevede l'uso dei coefficienti moltiplicativi nell'operazione di trasformazione ma ne tiene conto in fase di antitrasformazione: questo produce un semplice scalamento costante per tutti i coefficienti ottenuti. (3772)

TABELLA I. ESEMPI DI STRUTTURE SPAZIALI E DEI RELATIVI COEFFICIENTI DCT

MATRICE SEGNALE

16.00	235.00	16.00	235.00	16.00	235.00	16.00	235.00
16.00	235.00	16.00	235.00	16.00	235.00	16.00	235.00
16.00	235.00	16.00	235.00	16.00	235.00	16.00	235.00
16.00	235.00	16.00	235.00	16.00	235.00	16.00	235.00
16.00	235.00	16.00	235.00	16.00	235.00	16.00	235.00
16.00	235.00	16.00	235.00	16.00	235.00	16.00	235.00
16.00	235.00	16.00	235.00	16.00	235.00	16.00	235.00
16.00	235.00	16.00	235.00	16.00	235.00	16.00	235.00

MATRICE COSENO

8032.00	-893.16	0.00	-1053.56	0.00	-1576.76	0.00	-4490.23
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

PATTERN 7

MATRICE SEGNALE

16.00	16.00	16.00	16.00	16.00	16.00	16.00	16.00
235.00	235.00	235.00	235.00	235.00	235.00	235.00	235.00
16.00	16.00	16.00	16.00	16.00	16.00	16.00	16.00
235.00	235.00	235.00	235.00	235.00	235.00	235.00	235.00
16.00	16.00	16.00	16.00	16.00	16.00	16.00	16.00
235.00	235.00	235.00	235.00	235.00	235.00	235.00	235.00
16.00	16.00	16.00	16.00	16.00	16.00	16.00	16.00
235.00	235.00	235.00	235.00	235.00	235.00	235.00	235.00

MATRICE COSENO

8032.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
-893.16	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
-1053.56	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
-1576.76	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
-4490.23	0.00	0.00	0.00	0.00	0.00	0.00	0.00

PATTERN 8

MATRICE SEGNALE

16.00	235.00	16.00	235.00	16.00	235.00	16.00	235.00
235.00	16.00	235.00	16.00	235.00	16.00	235.00	16.00
16.00	235.00	16.00	235.00	16.00	235.00	16.00	235.00
235.00	16.00	235.00	16.00	235.00	16.00	235.00	16.00
16.00	235.00	16.00	235.00	16.00	235.00	16.00	235.00
235.00	16.00	235.00	16.00	235.00	16.00	235.00	16.00
16.00	235.00	16.00	235.00	16.00	235.00	16.00	235.00
235.00	16.00	235.00	16.00	235.00	16.00	235.00	16.00

MATRICE COSENO

8032.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	-113.83	0.00	-134.27	0.00	-200.96	0.00	-572.27
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	-134.27	0.00	-158.39	0.00	-237.04	0.00	-675.04
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	-200.96	0.00	-237.04	0.00	-354.78	0.00	-1010.27
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	-572.27	0.00	-675.04	0.00	-1010.27	0.00	-2877.02

PATTERN 9

MATRICE SEGNALE

235.00	213.00	191.00	169.00	147.00	124.00	102.00	80.00
213.00	204.00	186.00	165.00	144.00	122.00	100.00	79.00
191.00	186.00	172.00	155.00	136.00	116.00	95.00	74.00
169.00	165.00	155.00	141.00	124.00	106.00	87.00	67.00
147.00	144.00	136.00	124.00	110.00	93.00	75.00	57.00
124.00	122.00	116.00	106.00	93.00	79.00	62.00	45.00
102.00	100.00	95.00	87.00	75.00	62.00	47.00	31.00
80.00	79.00	74.00	67.00	57.00	45.00	31.00	16.00

MATRICE COSENO

7294.00	1541.90	-198.50	113.52	-52.33	29.09	-17.28	6.05
1541.90	310.39	42.13	27.03	4.70	4.76	-0.12	3.22
-198.50	42.13	33.26	15.14	10.23	4.75	2.12	1.17
113.52	27.03	15.14	10.91	8.33	5.16	4.29	2.21
-52.33	4.70	10.23	8.33	7.00	3.86	1.18	2.38
29.09	4.76	4.75	5.16	3.86	6.02	1.18	0.06
-17.28	-0.12	2.12	4.29	1.18	1.18	0.74	-0.35
6.05	3.22	1.17	2.21	2.38	0.06	-0.35	0.69

PATTERN 10

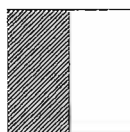
1327

segue

MATRICE SEGNALE

16.00	16.00	16.00	16.00	235.00	235.00	235.00	235.00
16.00	16.00	16.00	16.00	235.00	235.00	235.00	235.00
16.00	16.00	16.00	16.00	235.00	235.00	235.00	235.00
16.00	16.00	16.00	16.00	235.00	235.00	235.00	235.00
16.00	16.00	16.00	16.00	235.00	235.00	235.00	235.00
16.00	16.00	16.00	16.00	235.00	235.00	235.00	235.00
16.00	16.00	16.00	16.00	235.00	235.00	235.00	235.00
16.00	16.00	16.00	16.00	235.00	235.00	235.00	235.00

MATRICE COSENO

8032.00	-4490.23	0.00	1576.76	0.00	-1053.56	0.00	893.16
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

PATTERN 1

MATRICE SEGNALE

16.00	16.00	16.00	16.00	16.00	16.00	16.00	16.00
16.00	16.00	16.00	16.00	16.00	16.00	16.00	16.00
16.00	16.00	16.00	16.00	16.00	16.00	16.00	16.00
16.00	16.00	16.00	16.00	16.00	16.00	16.00	16.00
235.00	235.00	235.00	235.00	235.00	235.00	235.00	235.00
235.00	235.00	235.00	235.00	235.00	235.00	235.00	235.00
235.00	235.00	235.00	235.00	235.00	235.00	235.00	235.00
235.00	235.00	235.00	235.00	235.00	235.00	235.00	235.00

MATRICE COSENO

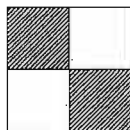
8032.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
-4490.23	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1576.76	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
-1053.56	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
893.16	0.00	0.00	0.00	0.00	0.00	0.00	0.00

PATTERN 2

MATRICE SEGNALE

16.00	16.00	16.00	16.00	235.00	235.00	235.00	235.00
16.00	16.00	16.00	16.00	235.00	235.00	235.00	235.00
16.00	16.00	16.00	16.00	235.00	235.00	235.00	235.00
16.00	16.00	16.00	16.00	235.00	235.00	235.00	235.00
235.00	235.00	235.00	235.00	16.00	16.00	16.00	16.00
235.00	235.00	235.00	235.00	16.00	16.00	16.00	16.00
235.00	235.00	235.00	235.00	16.00	16.00	16.00	16.00
235.00	235.00	235.00	235.00	16.00	16.00	16.00	16.00

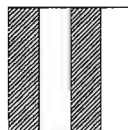
MATRICE COSENO

8032.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	-2877.02	0.00	1010.27	0.00	-675.04	0.00	572.27
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	1010.27	0.00	-354.76	0.00	237.04	0.00	-200.96
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	-675.04	0.00	237.04	0.00	-158.39	0.00	134.27
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	572.27	0.00	-200.96	0.00	134.27	0.00	-113.83

PATTERN 3

MATRICE SEGNALE

16.00	16.00	235.00	235.00	16.00	16.00	235.00	235.00
16.00	16.00	235.00	235.00	16.00	16.00	235.00	235.00
16.00	16.00	235.00	235.00	16.00	16.00	235.00	235.00
16.00	16.00	235.00	235.00	16.00	16.00	235.00	235.00
16.00	16.00	235.00	235.00	16.00	16.00	235.00	235.00
16.00	16.00	235.00	235.00	16.00	16.00	235.00	235.00
16.00	16.00	235.00	235.00	16.00	16.00	235.00	235.00
16.00	16.00	235.00	235.00	16.00	16.00	235.00	235.00

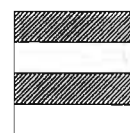
MATRICE COSENO

8032.00	-1859.91	0.00	-3806.63	0.00	2543.51	0.00	369.96
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

PATTERN 4

MATRICE SEGNALE

16.00	16.00	16.00	16.00	16.00	16.00	16.00	16.00
16.00	16.00	16.00	16.00	16.00	16.00	16.00	16.00
235.00	235.00	235.00	235.00	235.00	235.00	235.00	235.00
235.00	235.00	235.00	235.00	235.00	235.00	235.00	235.00
16.00	16.00	16.00	16.00	16.00	16.00	16.00	16.00
16.00	16.00	16.00	16.00	16.00	16.00	16.00	16.00
235.00	235.00	235.00	235.00	235.00	235.00	235.00	235.00
235.00	235.00	235.00	235.00	235.00	235.00	235.00	235.00

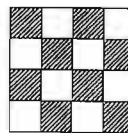
MATRICE COSENO

8032.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
-1859.91	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
-3806.63	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2543.51	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
369.96	0.00	0.00	0.00	0.00	0.00	0.00	0.00

PATTERN 5

MATRICE SEGNALE

16.00	16.00	235.00	235.00	16.00	16.00	235.00	235.00
16.00	16.00	235.00	235.00	16.00	16.00	235.00	235.00
235.00	235.00	16.00	16.00	235.00	235.00	16.00	16.00
235.00	235.00	16.00	16.00	235.00	235.00	16.00	16.00
16.00	16.00	235.00	235.00	16.00	16.00	235.00	235.00
16.00	16.00	235.00	235.00	16.00	16.00	235.00	235.00
235.00	235.00	16.00	16.00	235.00	235.00	16.00	16.00
235.00	235.00	16.00	16.00	235.00	235.00	16.00	16.00

MATRICE COSENO

8032.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	-493.62	0.00	-1010.27	0.00	675.04	0.00	98.19
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	-1010.27	0.00	-2067.70	0.00	1381.59	0.00	200.96
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	675.04	0.00	1381.59	0.00	-923.15	0.00	-134.27
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	98.19	0.00	200.96	0.00	-134.27	0.00	-19.53

PATTERN 6

BIBLIOGRAFIA

B.1. *Generali*

— CCIR, *Raccomandazione 601*, Volume XI, Parte 1, Ginevra 1982.

— BARBIERI G.: *Codifica di segnali televisivi ad alta definizione per la trasmissione numerica via satellite*, «Elettronica e Telecomunicazioni», n. 1, 1988.

— GONZALEZ R. C., WINTZ P.: *Digital Image Processing*, Addison-Wesley, 1987.

— HALL E. L.: *Computer Image Processing and Recognition*, Academic Press, 1979.

— WADE J. G.: *Signal Coding and Processing: An Introduction Based on Video Systems*, Ellis Horwood, 1987.

— WINTZ P.: *Transform Picture Coding*, «Proceedings of the IEEE», vol. 60, luglio 1972.

— MUSMANN H. G., PIRSCH P., GRALLERT H.: *Advances in Picture Coding*, «Proceedings of the IEEE», vol. 73, n. 4, aprile 1985.

— NETRAVALI A. N., LIMB J. O.: *Picture Coding: a Review*, «Proceedings of the IEEE», vol. 68, n. 3, marzo 1980.

B.2 *Riduzione della ridondanza*

— HUFFMAN D. A.: *A Method for the Construction of Minimum-Redundancy Codes*, «Proceedings of the IRE», settembre 1952.

— GILBERT E. N.: *Codes Based on Inaccurate Source Probabilities*, «IEEE Transactions on Information Theory», vol. IT-17, n. 3, maggio 1971.

— ANDREWS H. C.: *Entropy Considerations in the Frequency Domain*, «Proceedings of the IEEE», gennaio 1968.

— O'NEAL J. B.: *Bounds on Subjective Performance Measures for Source Encoding Systems*, «IEEE Transactions on Information Theory», vol. IT-17 n. 3, maggio 1971.

— GRANRATH D. J.: *The Role of Human Visual Models in Image Processing*, «Proceedings of the IEEE», vol. 69, n. 5, maggio 1981.

— TONGE G. J.: *The Sampling of Television Images*, IBA Experimental & Development, Report 112/81.

— TONGE G. J.: *Three-Dimensional Filters for Television Sampling*, IBA Experimental & Development, Report 117/82.

— SABATIER, KRETZ: *Sampling the Components of 625-Lines Colour Television Signal*, «EBU Review», n. 178, ottobre 1978.

— DUBOIS E.: *The Sampling and Reconstruction of Time-Varying Imagery with Application in Video Sistems*, «Proceedings of the IEEE», vol. 73, n. 4, aprile 1985.

B.3 *Codifica predittiva*

— CONNOR D. J., BRAINARD R. C., LIMB J. O.: *Intraframe Coding for Picture Transmission*, «Proceedings of the IEEE», vol. 60, n. 7, luglio 1972.

— ZETTERBERG L. H., ERICSSON S., BRUSEWITZ H.: *Interframe DPCM with Adaptive Quantization and Entropy Coding*, «IEEE Transactions on Communications», vol. COM-30, n. 8, agosto 1982.

— O'NEAL J. B.: *Differential Pulse-Code Modulation (PCM) with Entropy Coding*, «IEEE Transactions on Information Theory», vol. IT-22, n. 2, marzo 1976.

— PIRSCH P.: *Adaptive Intra-Interframe DPCM Coder*, «The Bell System Technical Journal», maggio-giugno 1982.

B.4 *Trasformazioni ortogonali*

— PRATT W. K., KANE J., ANDREWS H. C.: *Hadamard Transform Image Coding*, «Proceedings of the IEEE», vol. 57, n. 1, gennaio 1969.

— ALBERICO F., BOVE F.: *La trasformata di Hadamard nella trasmissione numerica di immagini*, «Elettronica e Telecomunicazioni», n. 3, 1973.

— PONCIN J.: *Utilisation de la transformation de Hadamard pour le codage et la compression de signaux d'images*, «Annales de Télécommunications», T. 26, n. 7-8, 1971.

— SCHWARTZ P.: *Analyse de la compression d'information sur la transformée de Fourier d'une image*, «Annales de Télécommunications», T. 26, n. 3-4, 1971.

— PEARLMAN W. A., GRAY R. M.: *Source Coding of the Discrete Fourier Transform*, «IEEE Transactions on Information Theory», vol. IT-24, n. 6, novembre 1978.

— NAKAGAWA M., MIYAHARA M.: *Generalized Karhunen-Loeve Transformation I (Theoretical Consideration)*, «IEEE Transactions on Communications», vol. COM-35, n. 2, febbraio 1987.

— MIYAHARA M., KOTANI K.: *Block Distortion in Orthogonal Transform Coding - Analysis, Minimization, and Distortion Measure*, «IEEE Transactions on Communications», vol. COM-33, n. 1, gennaio 1985.

B.5 *Trasformata coseno*

— AHMED N., NATARAJAN T., RAO K. R.: *Discrete Cosine Transform*, «IEEE Transactions on Computers», vol. C-23, gennaio 1974.

— MAKHOUL J.: *A Fast Cosine Transform in One and Two Dimensions*, «IEEE Transactions on Acoustics, Speech, and Signal Processing», vol. ASSP-28, n. 1, febbraio 1980.

— REININGER R. C., GIBSON J. D.: *Distributions of the Two Dimensional DCT Coefficients for Images*, «IEEE Transactions on Communications», vol. COM-31, n. 6, giugno 1983.

— NILL N. B.: *A Visual Model Weighted Cosine Transform for Image Compression and Quality Assessment*, «IEEE Transactions on Communications», vol. COM-33, n. 6, giugno 1985.

— CLARKE R. J.: *Spectral response of the Discrete Cosine and Walsh-Hadamard Transforms*, «IEE PROC.», vol. 130, Pt. F, n. 4, giugno 1983.

— WANG Z.: *On Computing the Discrete Fourier and Cosine Transforms*, «IEEE Transactions on Acoustics, Speech, and Signal Processing», vol. ASSP-33, n. 4, ottobre 1985.

— GI LEE B.: *A New Algorithm to Compute the Discrete Cosine Transform*, «IEEE Transactions on Acoustics, Speech, and Signal Processing», vol. ASSP-32, n. 6, dicembre 1984.

B.6 *Applicazioni DCT*

— CESA, MODENA, SICURANZA: *2D DCT Intrafield Coding on High-Quality TV Signals*, Conference on DSP, Firenze 1987.

— CUCCHI, MODENA: *Component Coding of TV Signal Using Intrafield DCT Transform: Simulation Results and Considerations about Hardware Complexity*, 5ª World Communication Forum, Geneve, ottobre 1987.

CODIFICA DEL SEGNALE TELEVISIVO NUMERICO:

SISTEMI DI RIDUZIONE DELLA RIDONDANZA MEDIANTE L'USO DELLA TRASFORMATA COSENO DISCRETA

M. BARBERO, M. STROPPIANA*

SOMMARIO — Presso il Centro Ricerche RAI vengono studiate le tecniche di riduzione della ridondanza del segnale video. La trasformata coseno discreta (DCT) e la codifica mediante codici a lunghezza variabile (VLC) si sono dimostrate estremamente efficienti. L'ottimizzazione di tali tecniche è effettuata mediante simulazioni per mezzo di calcolatori elettronici. Per al codifica numerica del segnale video sono normalmente usati 16 bit per ciascun elemento di immagine (pel). Un sistema basato sulla DCT intra-field che consente di ridurre, mediante, a circa 3 bit/pel il flusso di dati è stato sviluppato da un'industria italiana; esso è adatto alla codifica di segnali compositi (PAL e NTSC) e a componenti. L'uso della DCT ibrida, che sfrutta anche la ridondanza temporale, consentirà la realizzazione di co-decodificatori più efficienti: riducendo il numero di bit necessari a codificare ogni elemento di immagine fino ad un minimo di 1 bit/pel. Tali codec, nelle versioni adatte a segnali di definizione normale e ad alta definizione (HDTV), sono in avanzato stadio di sviluppo. Nell'articolo vengono descritti: gli algoritmi, le caratteristiche principali dei sistemi proposti, le tecniche che permetteranno ulteriori miglioramenti, e le principali applicazioni previste. È disponibile la versione in lingua inglese di questo articolo.

SUMMARY — ***Digital TV signal coding: systems for the redundancy reduction based on the Discrete Cosine Transform.*** *At the RAI Research Centre studies on the bit-rate reduction techniques for the video signal are carried on. The Discrete Cosine Transform (DCT) and the Variable Length Coding (VLC) demonstrated to be very efficient. The optimization of these techniques is performed by computer simulations. To code the video signal 16 bits are normally used for each picture element (pel). A system based on the intra-field DCT which allows the reduction to 3 bits per pel, on average, has been developed by an Italian manufacturer; it can be used both for the video composite signals (NTSC and PAL) and the component video signals. The adoption of the hybrid DCT, which makes use of the temporal redundancy, will allow the implementation of more efficient codecs: the reduction can be down to 1 bit per pel. These codecs, both for the standard definition and for the high definition TV, are in an advanced development stage. The article analyzes the following items: the algorithms, the main characteristics of the proposed systems, the techniques for further improvements, and the main applications. The English version of this article is available.*

1. Introduzione

In articoli precedenti si è introdotto l'argomento della codifica numerica del segnale televisivo (bibl 1) e sono stati analizzati i metodi di riduzione della ridondanza che consentono di codificare tale segnale con un flusso di informazione, o bit-rate, minimo, al fine di permetterne la trasmissione o la memorizzazione, compatibilmente con la capacità ed il costo degli attuali canali numerici (bibl. 2). Nell'ultimo articolo citato è descritto l'algoritmo della Trasformata Coseno Discreta (DCT, Discrete Cosine Transform), che presenta un'elevata efficienza e che per tale ragione è oggetto di studi ed è alla base dei codecodificatori in via di sviluppo in Italia. Si rimanda a tale articolo sia perché tutti gli argomenti in esso trattati sono qui considerati noti, sia perché è corredato di una ricca bibliografia.

In questo articolo si effettuerà una panoramica sulle ricerche che hanno dato origine alla realizzazione pratica di sistemi di trasmissione dei segnali video compositi (PAL, NTSC) e a componenti e che consentiranno di sviluppare sistemi per la televisione ad alta definizione (HDTV, High Definition TeleVision).

Dopo un breve cenno all'attuale normativa per quanto riguarda la codifica numerica del segnale televisivo e quella per l'uso dei canali di trasmissione numerica, si analizzano gli algoritmi di codifica basati su DCT. In pratica si segue un criterio cronologico nel descrivere le fasi successive della ricerca, tale criterio coincide con quello della crescente complessità realizzativa.

Si descrive la DCT bidimensionale puramente intrafield: essa opera su blocchi di elementi di immagine, ossia pel, appartenenti a singoli semiquadri. Si passa poi ad analizzare la DCT tridimensionale: essa opera opera su blocchi costituiti da pel appartenenti a quadri consecutivi. Infine si esamina la DCT ibrida: essa sfrutta, come quella tridimensionale, la ridondanza temporale presente

(*) Ing. Marzio Barbero, dott. Mario Stroppiana del Centro Ricerche RAI (Torino).
Dattiloscritto pervenuto alla Redazione l'11 aprile 1989.

nel segnale televisivo e, quando ciò consente una riduzione della quantità di informazione da trasmettere, codifica le differenze fra i pel appartenenti a semiquadri o quadri adiacenti.

Gli studi e la realizzazione pratica dei sistemi di codifica basati su DCT per la trasmissione punto a punto del segnale televisivo, sia quello a definizione normale, sia quello ad alta definizione, sono svolti nell'ambito del progetto europeo EUREKA EU256. A tale progetto collaborano, oltre alla RAI, le seguenti organizzazioni italiane e spagnole: Telettra SpA e Telettra Espana S.A., RTVE (radio televisione spagnola) e la Universidad Politecnica de Madrid.

2. Norme di codifica e di trasmissione

2.1 NORME DI CODIFICA DEL SEGNALE TELEVISIVO

Sono oggetto della Raccomandazione 601 del CCIR i parametri di codifica numerica per il segnale video a componenti, parametri riassunti in tabella 1 (bibl. 3). La Raccomandazione 656 riguarda invece le interfacce, parallelo e seriale, per collegare gli apparati di studio che operano in numerico (bibl. 4).

L'esistenza di una norma internazionale ha permesso di sviluppare apparati video professionali operanti con segnali numerici ed interfacciabili direttamente. È attualmente possibile utilizzare studi televisivi completamente numerici, con vantaggi sulla qualità del prodotto finale, soprattutto in caso di programmi complessi, che richiedono più operazioni successive sullo stesso segnale.

Attualmente in sede internazionale sono in discussione varie proposte di normalizzazione per il segnale HDTV.

Si prevede che, per HDTV, l'uso delle tecniche numeriche, almeno per quanto riguarda gli apparati professionali, sarà fin dall'inizio preponderante. Ad esempio, i primi esemplari di videoregistratori utilizzano tecniche analogiche, ma già nell'immediato futuro sono previsti quelli di tipo numerico.

Sarebbe auspicabile l'adozione di una norma unica mondiale, risolvendo così i problemi che si pongono attualmente per la TV a definizione normale; problemi di sviluppo di apparati diversi e di conversione di standard che comportano aumenti di costo e l'introduzione di degradamenti.

Vi sono però spinte a mantenere due formati, uno caratterizzato da una frequenza di trama a 60 Hz ed uno a 50 Hz, allo scopo di facilitare la conversione dei formati HDTV ai corrispondenti formati a definizione normale. D'altro canto si vorrebbe che i due formati fossero caratterizzati dal maggior numero possibile di parametri uguali.

Un primo criterio è quello di mantenere costante il bitrate necessario alla registrazione e trasmissione. Un'altro criterio è quello di mantenere costante il formato d'immagine, numero di pel per riga e numero di righe attive per quadro.

2.2 NORME SULLA CANALIZZAZIONE PER COLLEGAMENTI NUMERICI

Per trasferire il segnale fra due località distanti fra loro esistono due alternative: l'uso di nastro magnetico (il sistema di registrazione per il segnale video a componenti numerico è descritto in bibl. 5), oppure l'uso di reti di trasmissione. I canali di trasmissione per i segnali televisivi sono attualmente di tipo analogico, normalmente operanti con segnali compositi (PAL, NTSC, SECAM).

Come si è detto precedentemente, però, la tendenza attuale è quella di produrre e trattare i segnali televisivi sotto forma numerica e questa tendenza sarà ancora maggiore nel caso di HDTV.

Attualmente esistono delle gerarchie di multiplazione per la trasmissione di segnali numerici, elencate nella tabella 2. L'accesso alle reti numeriche avviene quindi in generale ad uno dei quattro livelli indicati in tabella (Racc. CCITT G.703).

In ambito CCITT sono in via di definizione i livelli di accesso alla rete integrata (ISDN, Integrated Services Digital Network) e la situazione attuale è riassunta in tabella 3.

Molti gestori di reti sono favorevoli all'introduzione di una rete a pacchetti di lunghezza fissa che consentano una trasmissione di tipo asincrono (ATM, Asynchronous

Tabella 1.
SPECIFICHE DELLO STANDARD VIDEO NUMERICO (RACC. CCIR 601)

parametri	sistemi	
linee per quadro semiquadri al secondo	525 60	625 50
segnali codificati	Y, C_B, C_R	
numero totale di campioni per linea — luminanza Y — crominanze C_B, C_R	 858 429	 864 432
struttura di campionamento	ortogonale, ripetitiva a livello di riga, semiquadro e quadro. I campioni CR e CB sono coposizionati con i campioni Y dispari (1°, 3° ...) di ciascuna riga.	
frequenza di campionamento — luminanza Y — crominanze CB, CR	 13,5 MHz 6,75 MHz	
forma di codifica campionamento	PCM a quantizzazione uniforme, 8 bit per campione per segnale di luminanza e ciascun segnale differenza di colore.	
numero di campioni per riga attiva — luminanza Y — crominanze CB, CR	 720 360	
corrispondenza fra i livelli del segnale analogico e i livelli di quantizzazione — luminanza — crominanze CB, CR	220 livelli quantizzaz. nero = livello 16 bianco = livello 235 225 livelli quantizzaz. zero = livello 128	
uso delle parole di codice	0 e 255 per sincronizzazione da 1 a 254 per video	

Tabella 2.

GERARCHIE DI TRASMISSIONE NUMERICA
(RACC. CCITT G.703)

livello	bit-rate [Mbit/s]		
1	1,544		2,048
2	6,312		8,448
3	32,064	44,736	34,368
4	97,728		139,264
usato in	Giappone	Stati Uniti	Europa

Nota: il primo livello è ottenuto dalla multiplazione sincrona di un certo numero di canali a 64 kbit/s, con informazioni aggiuntive per le commutazioni e la supervisione del sistema; 1544 kbit/s corrispondono a 24 canali da 64 kbit/s e 2048 kbit/s corrispondono a 30 canali. I livelli superiori sono ottenuti da multiplazione asincrona di quelli inferiori, con giustificazione, per sopperire a piccole variazioni del bit-rate nominale.

Transfer Mode). Tali sistemi sono particolarmente adatti per il traffico di dati generati in maniera discontinua, ad esempio fra elaboratori elettronici, ma potrebbero risultare più costosi e di più difficile impiego nel caso in cui i dati vengano generati in grande quantità ed in modo continuo, come nel caso del segnale video numerico.

In ogni caso si prevede che l'uso di sistemi di trasmissione numerica sia in futuro una scelta obbligata poiché consentono l'uso della stessa rete in modo flessibile, per molteplici servizi. Anche quei servizi, come quello televisivo, che fino ad oggi hanno utilizzato reti analogiche dovranno quindi adattarsi a questa tendenza verso la trasmissione numerica.

2.3 NORME RELATIVE A SISTEMI DI RIDUZIONE DELLA RIDONDANZA

Per i motivi precedentemente esposti occorre stabilire una norma di trasmissione dei segnali televisivi numerici. Poiché la capacità dei canali attualmente disponibili è limitata e, comunque, il costo dell'uso del canale è proporzionale alla capacità impiegata, si è deciso di studiare dei sistemi che consentano di minimizzare il bit-rate richiesto.

Il gruppo CMTT/2 ha il compito di definire un sistema di codifica per il segnale video conforme alla Racc. 601 adatto ad essere trasmesso in canali di capacità pari ai livelli H_{21} e H_{22} e che garantisca una *qualità contributo*. Si intende come qualità contributo quella che consente di effettuare certi tipi di trattamenti sul segnale video, ad esempio il «chroma-key» (intarsio di un'immagine, modificando lo sfondo) e lo «slow-motion» (rallentamento temporale).

Tabella 3.

GERARCHIE DI TRASMISSIONE NUMERICA
LIVELLI H PER ISDN

LIVELLO	BIT-RATE [MBIT/S]	
H_0	0,384	
H_1	1,536	1,920
H_2	43 – 44,776 (H_{22})	32,768 (H_{21})
H_3	60 – 70	
H_4	132 – 138,24	

Nota: alcuni dei valori non sono ancora stati fissati in modo definitivo, per cui sono indicati delle gamme. Il livello H4 dovrebbe essere comune a entrambe le gerarchie. Il livello H3 verrà studiato con priorità inferiore. Il livello H'! a 32,768 Mbit/s verrà adottato in Europa, quello H22 a circa 44 Mbit/s negli Stati Uniti.

Attualmente sono allo studio sistemi basati su tecniche predittive (DPCM, Differential Pulse Code Modulation) e tecniche utilizzanti la DCT.

Gli stessi algoritmi possono essere applicati ai segnali HDTV ed in questo caso ci si ripromette di ridurre il bit-rate in modo da utilizzare un canale di livello H_4.

In futuro si potranno impiegare gli stessi algoritmi, tecniche e architetture a sistemi di codifica caratterizzati da una *qualità distribuzione*, ovvero sufficiente per la trasmissione, eventualmente circolare, agli utenti finali. In questo caso i bit-rate richiesti sono almeno dimezzati rispetto ai precedenti.

Vi sono diversi gruppi di lavoro e progetti internazionali coinvolti nello studio di tali sistemi; oltre ai già citati CMTT/2 e EU256, alcuni gruppi specialistici UER (V1/RDB e V1/HDTV), RACE (HIVITIS), COST.

3. Sistema basato su DCT bidimensionale

3.1 PARAMETRI DI CODIFICA

L'algoritmo fondamentale alla base dei sistemi di ridondanza oggetto di studio presso il Centro Ricerche RAI è la DCT. In base a prove preliminari, ed in accordo con quanto riportato in letteratura (figura 1), un blocco di 8 · 8 pel è stato scelto come elemento fondamentale su cui applicare la trasformata bidimensionale. Tale valore è il miglior compromesso fra l'efficienza dell'algoritmo, che cresce al crescere delle dimensioni del blocco, e la complessità realizzativa, proporzionale al numero di operazioni necessarie a calcolare i coefficienti della trasformata.

Nel caso del segnale numerico conforme alla raccomandazione 601 si individuano, all'interno di ciascun semiquadro un certo numero di *strisce*, ciascuna costituita da 8 righe video. Il numero di strisce è pari a 36 nel caso del sistema a 625/50 e a 31 nel caso del sistema 525/60.

Di ciascuna striscia fanno parte 90 blocchi costituiti da campioni di luminanza Y, 45 blocchi costituiti da campioni di crominanza C_B e infine 45 blocchi dovuti a C_R. Ad ogni area di immagine costituita da 16 · 8 pel corrispondono due blocchi Y, uno C_B ed uno C_R.

La trasformata viene applicata a ciascun blocco, sia esso di luminanza oppure di crominanza, ed il risultato di tale operazione è costituito dall'insieme di 64 coefficienti che presentano una distribuzione statistica delle ampiezze non uniforme: i coefficienti prossimi a quello corrispondente alla componente continua hanno in generale ampiezze maggiori rispetto a quelli relativi ad armoniche di alta frequenza spaziale.

3.2 ELIMINAZIONE DELLA RIDONDANZA STATISTICA

Utilizzando un codice a lunghezza variabile (VLC, Variable Length Code) si può sfruttare la distribuzione non uniforme dei coefficienti per ridurre il numero medio di

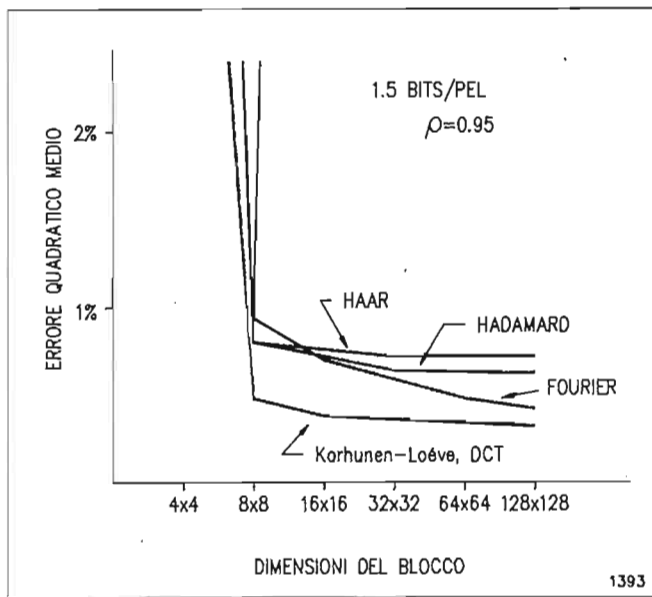



Fig. 1 — Prestazioni di differenti trasformate nel caso di una sorgente di Markov, immagine bidimensionale, caratterizzata da un fattore di correlazione orizzontale e verticale pari a 0,95 (tratto da NETRAVALI A. N.: *Picture Coding: a review*. «Proc. IEEE», vol. 68, n. 3, 1980, pag. 366). Le prestazioni sono espresse come errore quadratico medio in funzione delle dimensioni del blocco, in numero di pel, su cui è effettuata la trasformata.

bit assegnati a ciascun coefficiente. Fra i possibili codici a lunghezza variabile si è adottato il codice B_2 (bibl. 6). Questo codice ha una distribuzione delle lunghezze delle parole che si adatta molto bene alla distribuzione delle ampiezze dei coefficienti e presenta ottime caratteristiche dal punto di vista della realizzabilità circuitale.

Si è detto che i coefficienti di ampiezza bassa sono i più probabili, soprattutto per le alte frequenze spaziali. In particolare sono molto frequenti i coefficienti di valore zero. Scegliendo un'opportuna legge di scansione, e si è verificato che un percorso a zig-zag è conveniente, si fa sì che, statisticamente, i coefficienti nulli si trovino in posizioni contigue. È così possibile utilizzare la codifica di sequenze di zeri (run length encoding of zeroes). A ciascuna lunghezza di sequenze di zeri viene associata una parola VLC, come se fosse una possibile ampiezza dei coefficienti.

In particolare gli ultimi coefficienti del blocco, in base all'ordine di scansione, hanno una probabilità elevatissima di essere nulli. Conviene quindi, anziché codificarli con una parola VLC, anticipare la fine del blocco. Si utilizza a questo scopo una particolare parola VLC di fine blocco (EOB, End Of Block).

Lo studio statistico della distribuzione dei coefficienti, e conseguentemente i criteri che hanno portato alla scelta della legge di scansione e del codice a lunghezza variabile saranno oggetto di un articolo di prossima pubblicazione.

3.3 Quantizzazione adattativa

Il passaggio dal dominio temporale a quello della trasformata è un'operazione reversibile. A parte gli eventuali errori introdotti nel corso delle operazioni matematiche e dovute al fatto che si opera con parole binarie di lunghezza finita, mediante l'operazione di antitrasformata è sempre possibile riprodurre perfettamente il segnale originale.

Analogamente le operazioni di riduzione della ridondanza statistica sono reversibili.

Fino a questo punto, dunque, si è ottenuto, mediamente, una riduzione del numero di bit necessario a descrivere un'immagine senza introdurre nessun degradamento. Purtroppo l'entità della riduzione ottenuta non è fisso, ma dipende dal contenuto dell'immagine stessa. Non è quindi possibile limitarsi all'uso dei suddetti metodi quando si deve generare un flusso d'informazione costante, indipendente dalle caratteristiche delle immagini. È proprio questo il caso, in generale, della trasmissione e della registrazione, casi in cui la capacità del canale è fissa e limitata.

La soluzione adottata in questi casi è quella di eliminare l'informazione irrilevante, quella cioè poco importante dal punto di vista soggettivo.

L'occhio è in genere meno sensibile a distorsioni che nascono quando si utilizza una minor precisione per i coefficienti relativi alle frequenze spaziali più elevate. Ciò accade se si trasmette per tali coefficienti un minor numero di bit, tralasciando i meno significativi.

La quantizzazione non lineare dei coefficienti, diminuendo la precisione al crescere dell'ampiezza può avere lo scopo di ridurre, a parità di numero di livelli di quantizzazione, il rapporto segnale/ rumore di quantizzazione complessivo, poiché i valori di ampiezza maggiore sono meno probabili. Questa operazione, però, modifica la distribuzione statistica delle ampiezze dei coefficienti, rendendola più uniforme, per cui rende meno efficiente la codifica VLC. Si è infatti verificato che nel caso in cui il quantizzatore sia seguito da un codificatore VLC, la legge di quantizzazione lineare è quella ottimale.

La soluzione che è stata infine adottata è quella di ridurre la precisione con cui vengono trasmessi i coefficienti, dividendoli per un fattore di scalamento che dipende dalla differenza fra la quantità di dati prodotti dal codificatore e quelli prelevati dal canale, e trasmettendo solo i bit più significativi. In questo modo l'ampiezza dei coefficienti da codificare mediante VLC viene mediamente ridotta, permettendo l'uso di parole più corte, e soprattutto viene aumentato il numero di coefficienti nulli consentendo un maggiore uso della codifica di sequenze di zeri.

Poiché la quantità di dati prodotti dipende dal contenuto di informazione medio dell'immagine (entropia) e varia fortemente all'interno di un'immagine in funzione del dettaglio spaziale, è opportuno regolare la produzione di dati mediandola nel tempo: il valore dei fattori di scalamento è quindi controllato in base al riempimento di una memoria tampone (buffer di trasmissione).

3.4 Schema a blocchi del codificatore e del decodificatore

In figura 2 sono rappresentati gli schema a blocchi del codificatore e del decodificatore basati sulla DCT bidimensionale (intra-field).

Il segnale video nelle sue tre componenti Y, C_B e C_R

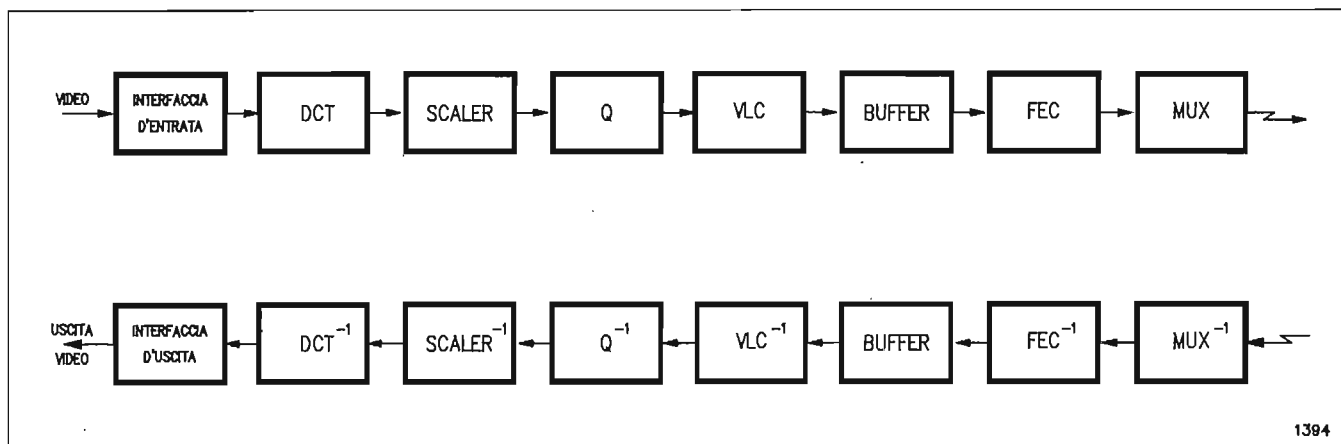



Fig. 2 — Schema a blocchi del codificatore e del decodificatore DCT intra-field. Il segnale d'ingresso è organizzato in blocchi di 8 · 8 pel dall'interfaccia d'entrata, viene effettuata la trasformata (DCT), i coefficienti vengono divisi per un fattore di scalamento (SCALER), quantizzati (Q), codificati con parole a lunghezza variabile (VLC) e memorizzati nella memoria di trasmissione (BUFFER). Dopo l'aggiunta del codice per la protezione dagli errori (FEC, Forward Error Correction) i dati vengono multiplati (MUX, Multiplexer) e trasmessi. Il decodificatore contiene i blocchi funzionali che effettuano le operazioni inverse.

viene organizzato in blocchi 8 · 8 dall'interfaccia d'entrata.

Il blocco di trasformazione effettua i calcoli e fornisce in uscita i blocchi di 64 coefficienti.

In base al riempimento del buffer di trasmissione vengono determinati i valori dei fattori di scalamento.

Segue il quantizzatore che, in questo caso, è lineare.

Vengono quindi determinate le lunghezze delle sequenze di zeri, generate le parole di fine blocco EOB e assegnate le parole VLC che vengono trasferite al buffer di trasmissione.

I dati devono essere protetti da un codice per consentire la correzione degli errori introdotti dal canale.

Infine i dati vengono organizzati in una trama di trasmissione opportuna e multiplati con gli altri servizi (audio, dati teletext, ecc...).

Lo schema a blocchi del decodificatore è perfettamente analogo a quello del codificatore e le funzioni svolte dai vari blocchi sono duali.

Fig. 3 — Scheda relativa all'unità di decodifica contenente il circuito integrato per il calcolo della DCT. Esso è in tecnologia HCMOS ed una complessità pari a 51 000 porte. Sono visibili due esemplari di tale circuito integrato: uno è usato per la componente di luminanza e uno per le due componenti di crominanza. Analogamente due esemplari sono usati nella scheda del codificatore.

3.5 Realizzazione pratica del codec

È stato realizzato in Italia un codificatore basato sui principi precedentemente descritti (bibl. 7).

Grazie all'uso di circuiti ad elevato fattore in integrazione (VLSI, Very Large Scale Integration) è stato possibile produrre un apparato di dimensioni e consumi contenuti.

Esso è in grado di operare su segnali numerici conformi alla raccomandazione 601. Esistono le versioni per 625 e 525 righe, adatte per le gerarchie CCITT a 34 e a 44 Mbit/s. Sono state prodotte anche versioni per segnali PAL e NTSC, nelle quali l'interfaccia d'entrata provvede a effettuare, dopo la conversione analogica-numerica, una decodifica in tre pseudo-componenti aventi una distribuzione statistica analoga a quella dei segnali Y, C_B e C_R.

I coefficienti di scalamento utilizzati sono del tipo $(\sqrt{2})^n$, vengono stabiliti striscia per striscia ogni 8 righe video), sono validi per tutti i blocchi di luminanza e crominanza appartenenti alla striscia; per ciascuna striscia viene inviato al decodificatore il fattore di scalamento applicato.

Il codice VLC utilizzato è, come si è già accennato, il B_2.

Per la protezione degli errori sono utilizzati codici a blocchi BCH; in particolare sono disponibili, in opzione, il BCH(255,239) e il BCH(511,493). Essendo caratterizzati da una distanza di Hamming pari a 5, tali codici permettono la correzione di errori semplici e doppi (bibl. 8 e Appendice di bibl. 9).

Per la realizzazione del codec sono stati realizzati sei

Fig. 4 — Codificatore e decodificatore basati su DCT intra-field: le dimensioni esterne sono 600 · 120 · 180 mm ed il consumo è pari a circa 30 W per entrambi gli apparati. Essi contengono, oltre alle quattro schede che realizzano la compressione dei dati video, le eventuali interfacce analogico/numerico per il segnale video e per quelli audio (fino ad un massimo di quattro) e l'alimentatore.

circuiti «semicustom»; è stata impiegata una tecnologia HCMOS da 1,5 μm e 1,25 μm. In particolare il «gate array» che realizza la funzione DCT è visibile in figura 3.

Il codificatore video (figura 4) è composto essenzialmente di quattro schede: interfaccia d'entrata (numerica 601, PAL o NTSC), scheda DCT, scheda di processamento e buffer, multiplatore e interfaccia di trasmissione. Il decodificatore ha una struttura analoga.

3.6 Prestazioni del sistema

Il sistema basato sulla DCT bidimensionale ha prestazioni buone per la maggior parte delle immagini. Purtroppo nel caso di immagini particolarmente critiche (molto dettagliate e quindi con scarsa correlazione spaziale) appaiono difetti di codifica individuabili come aumento e strutturazione del rumore, perdita di definizione.

Particolarmente significative sono le immagini in figura 5, ottenute in uscita dal co-decodificatore a 34 Mbit/s.

La prima non presenta difetti poiché il sistema riesce a ridurre il bit-rate sfruttando solamente, o principalmente, la ridondanza statistica presente nell'immagine.

Nella seconda immagine, invece, una porzione significativa è costituita da un segnale non correlato (rumore bianco, di luminanza e crominanza). Il sistema non può quindi eliminare la ridondanza statistica, che non esiste, ed è costretto, al fine di ridurre il bit-rate, a quantizzare in modo grossolano i coefficienti DCT. Ciò causa una strutturazione del rumore stesso, maggiormente visibile nella parte inferiore dell'immagine, quando il buffer tende ad essere sovraccarico, e meno visibile nella parte superiore dell'immagine, poiché durante l'intervallo di cancellazione di semiquadro il riempimento del buffer è diminuito. Si osservi come anche la porzione di immagine correlata subisca un degradamento dovuto alla quantizzazione grossolana.

Fortunatamente, come si è detto, immagini con caratteristiche analoghe alla prima sono molto probabili, mentre quelle simili alla seconda sono scarsamente probabili.

In pratica la seconda immagine è costruita ad-hoc per valutare il limite del sistema di codifica ed è particolarmente significativa per mettere in evidenza la struttura dei suoi difetti.

In ogni caso questo sistema è tecnologicamente all'avanguardia e ha caratteristiche migliori rispetto ad altri presenti sul mercato e basati su tecniche predittive (DPCM).

Un miglioramento sensibile delle prestazioni, ovviamente non nel caso di immagini totalmente incorrelate, si può ottenere sfruttando oltre alla ridondanza spaziale, mediante la DCT bidimensionale, anche la ridondanza temporale. Ciò può essere ottenuto con l'uso di DCT tridimensionale oppure di sistemi ibridi, DCT bidimensionale più tecniche predittive.

4. Sistemi che riducono la ridondanza temporale

4.1 Sistema basato su DCT tridimensionale

Nel caso di sistema utilizzante la DCT tridimensionale si è scelto, come elemento fondamentale su cui applicare la trasformata, il blocco costituito da 8 · 8 · 4 pel. Tale dimensione è l'estensione della struttura del blocco usato nella DCT bidimensionale. Il numero di piani temporali è stato scelto pari a quattro in base ad un compromesso tra efficienza dell'algoritmo e complessità realizzativa. Al crescere della profondità del blocco è possibile sfruttare maggiormente la ridondanza temporale del segnale video; infatti è evidente che, nel caso di immagini statiche, i coefficienti relativi alle frequenze temporali hanno ampiezze trascurabili, al limite nulli se fosse possibile eliminare i contributi dovuti al rumore e presenti nel caso delle sorgenti più comuni.

I piani temporali di ogni blocco possono essere quelli omologhi ottenuti da quattro quadri successivi oppure es-

Fig. 5 — Due semiquadri ottenuti a partire dalla sequenza « Diva with noise » (realizzata da CBC, Canada e ITV, Gran Bretagna) co-decodificata mediante il sistema DCT intra-field. Si evidenzia il comportamento del codec nel caso di immagine (la prima) con forte correlazione spaziale e nel caso di immagine (la seconda) caratterizzata da bassissima ridondanza (in prevalenza rumore gaussiano bianco). Nella seconda immagine sono visibili alcuni difetti dovuti al fatto che il codificatore è costretto ad eliminare informazione che non è irrilevante dal punto di vista soggettivo: all'interno del cerchio compare la struttura dei blocchi di 8·8 pel; la struttura del rumore codificato è diversa nella parte superiore dell'immagine, quando il modo di codifica è migliore, rispetto a quella nella parte inferiore, quando il buffer di trasmissione è in condizioni di riempimento eccessivo.

sere quattro semiquadri consecutivi. Nel primo caso si ha la massima efficienza dell'algoritmo, specialmente per quanto riguarda le zone d'immagine prive di movimento, ma si richiede una memoria pari a quattro quadri sia nel codificatore che nel decodificatore. La seconda soluzione è caratterizzata da una minore efficienza in quanto, a causa dell'interlacciamento di riga fra i semiquadri adiacenti, anche quelle parti d'immagine senza movimento generano coefficienti DCT temporali; essa presenta però il vantaggio di richiedere una memoria pari a solo due quadri.

Malgrado si possa prevedere che i costi delle memorie siano destinati a ridursi nel futuro e quindi la prima soluzione sia di un certo interesse, si è preferito limitare la valutazione alla seconda soluzione, quella in cui il blocco è costituito nella direzione temporale da quattro semiquadri consecutivi.

Anche nel caso di DCT tridimensionale, come in quello bidimensionale, si opera separatamente sui blocchi di luminanza e su quelli relativi alle due componenti di crominanza. Il calcolo della DCT tridimensionale avviene eseguendo la trasformazione monodimensionale lungo i tre assi del blocco da trasformare e, quindi, la sua formula è facilmente ricavabile da quella monodimensionale descritta in bibl. 2.

A partire da ciascun blocco si ottengono 256 coeffi-

cienti che presentano una distribuzione statistica delle ampiezze analoga a quella che si ottiene nel caso della trasformata bidimensionale: in particolare l'energia si concentra nella zona dei coefficienti corrispondenti alle frequenze, spaziali e temporali, più basse. È quindi possibile eliminare la ridondanza statistica utilizzando codici a lunghezza variabile e codificando con parole di codice anche le sequenze di coefficienti nulli.

La codifica delle sequenze di zeri comporta la determinazione di un percorso di scansione ottimo dei coefficienti all'interno del blocco in modo da massimizzarne la lunghezza. Un possibile percorso di scansione è quello ricavabile mediante una semplice estensione di quello usato nel caso bidimensionale e che procede a zig-zag lungo le tre dimensioni partendo dal vertice corrispondente al coefficiente relativo alla componente continua.

Tale percorso non è di fatto il più efficiente in quanto si osserva che, per la maggior parte delle immagini, le ampiezze dei coefficienti DCT diminuiscono molto più velocemente passando da un piano temporale al successivo di quanto diminuiscano all'interno del medesimo piano temporale.

In base alle prove effettuate su un certo numero di sequenze, ritenute rappresentative, si è determinato che la scansione «piano dopo piano» è quella più efficiente: si effettua la scansione a zig-zag all'interno del primo piano temporale e si passa via via a scandire nello stesso modo i piani successivi; in pratica ciascun piano viene scandito solo al termine della scansione dei piani precedenti.

Lo schema a blocchi del sistema che utilizza la DCT tridimensionale è del tutto analogo a quello della DCT bidimensionale. Le differenze tra i due tipi di codifica riguardano, infatti, solo la dimensione e la struttura dei blocchi costituiti dai campioni del segnale video, e ciò ha conseguenze sulla complessità circuitale ma non sulla organizzazione del sistema di co-decodifica.

4.2 SISTEMA IBRIDO

Per sfruttare la ridondanza temporale del segnale video si possono codificare mediante la trasformata le differenze tra il valore degli elementi di immagine ed un valore di predizione. Se il valore di predizione è sufficientemente prossimo a quello effettivo i valori delle differenze sono molto piccoli, al limite sono nulli, ed in questo caso anche i coefficienti della trasformata applicata alle differenze hanno un'ampiezza piccola, adatta per essere codificata con parole a lunghezza variabile. Poiché questo metodo di riduzione della ridondanza fa uso sia di tecniche di trasformazione che di tecniche di codifica predittiva (DPCM = Differential Pulse Code Modulation), esso viene indicato come sistema DCT ibrido.

Anche in questo caso le dimensioni del blocco su cui applicare la trasformata sono le stesse adottate per il sistema puramente intra-field, cioè 8 · 8 valori.

Si possono ottenere predizioni buone soprattutto per le parti d'immagine statiche o quasi statiche.

Nel caso di parti statiche il valore ottimale di predizione per ciascun pel è costituito dal pel corrispondente appartenente al quadro precedente (predizione inter-frame).

Nel caso in cui siano presenti lievi movimenti, il valore di predizione ottimo può non essere quello ottenibile dai pel distanti temporalmente 1/25 o 1/30 di secondo, bensì quello ottenibile a partire dai pel distanti 1/50 o 1/60 di secondo, appartenenti cioè al semiquadro precedente (predizione inter-field). A causa dell'interlacciamento di riga, però, il valore di predizione deve essere ricavato mediante interpolazione a partire dai campioni appartenenti alle due righe adiacenti del semiquadro precedente. Nel caso di elevato dettaglio verticale l'operazione di interpolazione può non fornire un valore di predizione sufficientemente buono.

Pertanto la scelta di utilizzare la ridondanza temporale e, in caso affermativo, di usare la predizione inter-frame o inter-field viene operata, per ciascun blocco 8 · 8 da codificare in base al minimo bit-rate risultate. In tabella 4 si riportano i guadagni ottenuti passando dalla predizione inter-frame a quella inter-field ed all'insieme delle due predizioni, quella inter-frame e quella inter-field.

La scelta ottimale tra i tre modi di codifica (intra-field, inter-field e inter-frame) dovrebbe essere effettuata, per ciascun blocco, «a posteriori», in base al numero minimo di bit necessari a codificare il blocco stesso. Tale tipo di scelta richiede però la trasformazione, la quantizzazione e la codifica di tre blocchi, uno per ciascun modo.

Per ridurre la complessità circuitale si è deciso di effettuare una scelta subottimale, ma «a priori», ossia la scelta del modo di codifica viene effettuata prima dell'operazione di trasformata. Si è verificato che una scelta basata sulla minima energia del blocco è quasi coincidente con quella «a posteriori». L'energia è calcolata a partire dai valori dei pel o da quelli ottenuti come differenza con i valori di predizione, ma escludendo l'energia associata alla componente continua del blocco.

4.3 CONFRONTO FRA I SISTEMI DCT TRIDIMENSIONALE E IBRIDO

Oltre al sistema ibrido precedentemente descritto, in cui la predizione è effettuata sugli elementi di immagine,

Tabella 4.
COMPARAZIONE DELLE PRESTAZIONI IN FUNZIONE DEI MODI INTER UTILIZZATI

seq.	oltre al modo intra-field	blocchi in [%] intra Y	intra C	i-fild Y	i-fild C	i-frm Y	i-frm C	S/N [dB] Y	S/N [dB] C
A	inter-frame	48	70			52	30	34,1	38,3
	inter-field	34	50	66	50			34,6	38,5
	entrambi	37	41	29	40	34	19	34,9	38,7
B	inter-frame	26	67			74	33	30,2	33,7
	inter-field	17	39	83	61			31,7	34,7
	entrambi	6	28	64	54	30	18	32,0	34,8

Nota: Per due sequenze (A e B) ed a un bit-rate di 20 Mbit/s sono stati determinati i valori percentuali corrispondenti al numero di blocchi di luminanza (Y) e crominanza (C) che vengono codificati in intra-field, inter-field ed inter-frame nei casi in cui, oltre al modo intra-field, siano usati: il modo inter-frame, il modo inter-field ed entrambi i modi inter-field ed inter-frame. Sono inoltre indicati i valori di rapporto segnale rumore (S/N) ottenuti nei vari casi per luminanza e crominanza. I valori dei rapporti segnale/rumore (S/N) espressi in dB sono definiti come $220/\sigma$ per la luminanza e $225/\sigma$ per la crominanza, dove σ indica il valore efficace del rumore; 220 è il numero di livelli corrispondente al salto bianco-nero della luminanza e 225 quello corrispondente al picco-picco di crominanza (Racc. 601). Le sequenze usate sono «Renata» (A) e «Renata Butterfly» (B), entrambe riprese presso il Centro Ricerche RAI.

Tabella 5.
CONFRONTO FRA I SISTEMI
DCT IBRIDO E DCT TRIDIMENSIONALE

seq	bit-rate [Mbit/s]	H-DCT S/N [dB] Y	H-DCT S/N [dB] C	3D-DCTN S/N [dB] Y	3D-DCTN S/N [dB] C
A	10	35,5	36,6	35,5	36,3
	15	38,6	37,8	38,5	38,0
	25	42,5	40.9	41,3	42,1
B	10	32,9	32.9	32,5	32,4
	15	36,1	34,8	35,5	34,2
	25	40,0	37,7	39,6	37,1

Nota: Confronto fra le prestazioni, espresse sotto forma di rapporto segnale rumore S/N per luminanza (Y) e crominanza (C), della trasformata di tipo ibrido (H-DCT) e di quella tridimensionale (3D-DCT). Sono state usate per le simulazioni a bit-rate diversi (pari a 10, 15, 25 Mbit/s) due sequenze «Doll» (A) e «Girls with inset» (B) entrambe riprese presso il CCETT, Francia.

è possibile sviluppare un sistema in cui la predizione è effettuata sui coefficienti DCT. In questo caso viene trasmessa, quando ciò conviene in termini di bit-rate, la differenza fra ciascun coefficiente ed il suo omologo relativo al quadro precedente. Questa seconda soluzione è risultata però meno efficiente di quella precedentemente descritta.

La scelta fra uso della DCT tridimensionale e uso di quella ibrida è stata effettuata in base ai seguenti criteri: efficienza, possibilità di ulteriori miglioramenti grazie all'ottimizzazione ed all'uso di tecniche di riduzione della ridondanza aggiuntive, ed infine complessità realizzativa.

È possibile comparare le prestazioni offerte dai due approcci analizzando i risultati, riportati in tabella 5, ottenuti mediante simulazione su due sequenze a diversi bit-rate. Il sistema ibrido ha prestazioni leggermente migliori di quello utilizzante la trasformata tridimensionale, soprattutto nel caso della seconda sequenza.

Per il sistema ibrido è inoltre possibile introdurre ulteriori miglioramenti: ad esempio ottimizzazione della predizione inter-field e tecniche di compensazione del movimento; tali miglioramenti non sono applicabili al sistema DCT tridimensionale. Per quanto riguarda la complessità, il sistema basato su DCT tridimensionale è concettualmente più semplice, essendo un'estensione di quello puramente spaziale, e quindi si presta ad una realizzazione circuitale più modulare. D'altro canto il sistema ibrido richiede l'uso di una minore quantità di memoria.

La scelta finale, a favore del sistema ibrido, è basata soprattutto su considerazioni di maggior possibilità di ottimizzazione e quindi, in ultima analisi, di migliori prestazioni ottenibili.

5. **Realizzazione del sistema DCT di tipo ibrido**

5.1 SCHEMA A BLOCCHI DEL CODIFICATORE

In figura 6 è rappresentato lo schema a blocchi del codificatore DCT ibrido.

Si osservi che alcuni blocchi sono identici a quelli presenti nel codificatore DCT bidimensionale di figura 2.

In particolare sono funzionalmente identici il blocco di trasformazione degli 8 · 8 elementi d'immagine, lo sca-

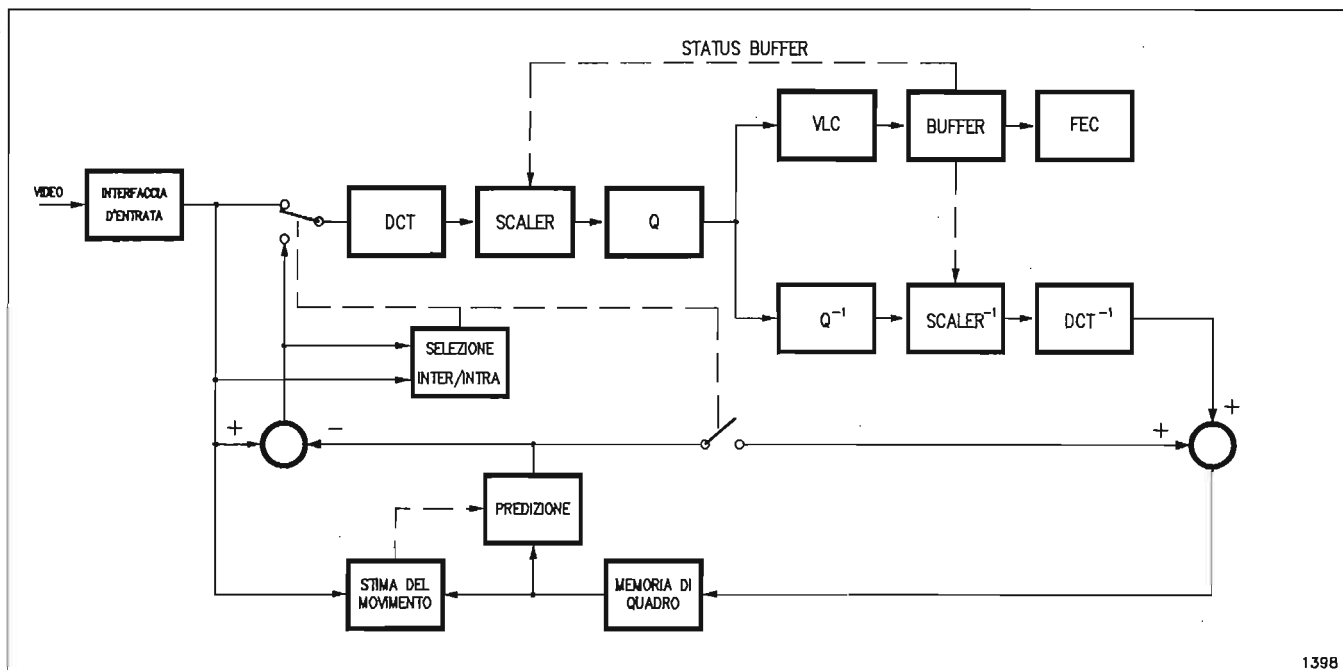



Fig. 6 — Schema a blocchi del codificatore DCT ibrido. I blocchi di interfaccia d'entrata, DCT, SCALER, Q, VLC, BUFFER, FEC sono analoghi a quelli descritti in figura 2 nel caso di codificatore DCT intra-field. In questo caso, però l'operazione di trasformata viene effettuata, in base alla decisione del blocco di selezione, sugli elementi d'immagine oppure sulle differenze fra questi ed i valori di predizione. Questi sono ottenuti a partire dai campioni appartenenti ai due semiquadri precedenti, ricostruiti per mezzo dei blocchi che effettuano le operazioni inverse di trasformazione, scalamento e quantizzazione e memorizzati nella memoria di quadro.

lamento dei coefficienti (del tipo $2^{n/8}$), il quantizzatore, l'assegnazione delle parole a lunghezza variabile ai coefficienti ed alle sequenze di zeri, il buffer e l'aggiunta della ridondanza per la protezione dagli errori.

Il blocco di 8 · 8 elementi da trasformare può essere, come nel caso del sistema DCT puramente intra-field, ottenuto dal semiquadro attuale oppure può essere ricavato come differenza fra i pel del semiquadro attuale e i valori di predizione ottenuti a partire dal quadro precedente.

Sono inoltre presenti: una memoria contenente i pel dei due semiquadri precedenti, un modulo di predizione che fornisce i valori interpolati in base al semiquadro immediatamente precedente o i pel coposizionati appartenenti al semiquadro ancora precedente, ed un modulo che effettua la selezione opportuna tra intra-field, inter-field e inter-frame in base al minimo dell'energia associata al blocco 8 · 8, escludendo la componente continua.

I valori contenuti nella memoria di quadro devono essere uguali, a meno di differenze dovute ad errori di trasmissione, a quelli disponibili dal lato decodificatore; a questo scopo sono presenti anche nel codificatore i blocchi funzionali di quantizzazione inversa, moltiplicazione per il fattore di scalamento e trasformazione inversa.

È possibile, ma non necessario, introdurre anche un blocco per la valutazione e, in decodifica, la compensazione del movimento. Utilizzando tale tecnica è possibile compensare i movimenti dell'intera immagine, dovuti, ad esempio, al movimento di traslazione della telecamera o «panning», oppure di porzioni d'immagine, facendo sì che il sistema operi in modo inter-frame anziché intra-field e inter-field e sfrutti maggiormente la correlazione temporale. Questa tecnica e i vantaggi che comporta saranno oggetto di un articolo di prossima pubblicazione.

5.2 PROTEZIONE DAGLI ERRORI DI CANALE E STRUTTURA DI TRAMA

Anche nel caso del sistema ibrido occorre proteggere i dati dagli errori introdotti dal canale e per far ciò si adotta lo stesso codice disponibile nel sistema puramente bidimensionale, ovvero il BCH(255,239).

Gli errori sui coefficienti che non vengono corretti danno origine, dopo l'operazione di antitrasformata, a gruppi di 8 · 8 o 16 · 8 pel incorrelati con il resto dell'immagine e quindi, in generale, molto visibili. Nel caso di sistema ibrido, a differenza del sistema puramente intra-field, l'effetto degli errori si propaga fino a quando i blocchi coposizionati nei semiquadri successivi non vengono trasmessi in modo intra-field.

È quindi necessario determinare una strategia di rinfresco o «refresh» che consiste nell'usare periodicamente il modo intra-field anche quando, in base al solo criterio di efficienza, i modi inter sarebbero scelti per lunghi intervalli di tempo.

La presenza, dal lato decodifica, della memoria di quadro consente di attuare, oltre alla correzione degli errori, anche la tecnica del mascheramento o «concealment» degli errori. Questa tecnica è usata anche in altri campi per esempio nel caso di trasmissione numerica del segnale sonoro (bibl. 9).

I problemi connessi alla protezione degli errori, oltre a quelli di recupero dei sincronismi di parola, di blocco e di trama video, hanno portato alla definizione di una struttura di trama organizzata a pacchetti di lunghezza fissa particolarmente efficiente e flessibile. Questi argomenti sono oggetto di un articolo di prossima pubblicazione.

5.3 PRESTAZIONI DEL SISTEMA

Il sistema di codifica è stato ottimizzato e valutato effettuando simulazioni mediante calcolatore su varie sequenze di immagini, più critiche della media.

In base a tali simulazioni, così come si è già verificato nel caso del sistema puramente intra-field, è possibile prevedere con ottima approssimazione quali saranno le prestazioni reali degli apparati prima di iniziarne la progettazione. I risultati che vengono quindi qui riportati dovrebbero corrispondere a quelli reali, ottenibili al completamento dei codec.

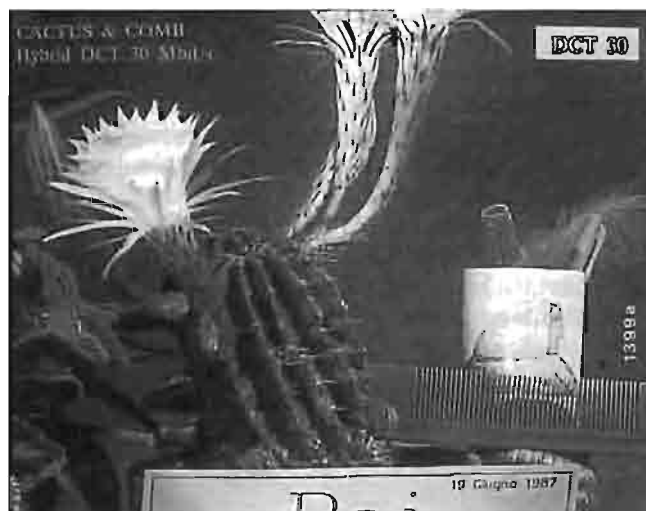



Fig. 7 — Semiquadro tratto dalla sequenza «Cactus and Comb» (ripresa presso il Centro Ricerche RAI) co-decodificata mediante sistema DCT ibrido. Le due immagini si riferiscono alla stessa sequenza prima e dopo l'effettuazione dell'operazione di Chroma-key (ottenuta dal segnale RGB analogico). Dimostrano che operazioni di post-produzione sono fattibili su segnali trasmessi o registrati con bit-rate dell'ordine di 30 Mbit/s.

Gli standard numerici di codifica a componenti prevedono 16 bit per ciascun pel (8 bit per la luminanza e 8 bit per, alternativamente, una delle due componenti di crominanza).

Se si escludono le sequenze appositamente prodotte per mettere in luce i problemi di codifica, si può prevedere che la qualità ottenibile con 3 bit per pel soddisfi i criteri imposti per la qualità contributo. In questo caso occorre ottimizzare la qualità oggettiva del segnale co-decodificato, massimizzando il rapporto segnale rumore. Infatti alcune operazioni di post-produzione, come il chroma-key, amplificano il rumore iniziale, rendendolo visibile.

In figura 7 è riprodotto il chroma-key ottenuto a partire da una sequenza conforme alla Racc. 601 co-decodificata a 30 Mbit/s.

Lo schema di codifica ibrido, ed in pratica il codec stesso, possono essere utilizzati a bit-rate inferiori. In figura 8 sono riprodotti: un quadro della sequenza «Renata», e le differenze fra lo stesso quadro trattato a diversi bit-rate (corrispondenti a 3, 1,5 e 1 bit/pel) e il quadro originale. In pratica viene così evidenziata la struttura del rumore di quantizzazione introdotto e il suo aumento al decrescere del numero medio di bit per pel. Si visualizzano le differenze perché in una comparazione diretta delle immagini i degradamenti sono scarsamente visibili, soprattutto se riprodotte fotograficamente.

Nel caso in cui l'obiettivo sia la distribuzione del segnale, e quindi si voglia migliorare soprattutto la qualità soggettiva, possono essere utilizzate le tecniche che considerano la curva di sensibilità dell'occhio, minore alle alte frequenze che alle basse, e che privilegiano una riduzione della risoluzione rispetto ad un aumento del rumore.

5.4 SISTEMI DI TRASMISSIONE PER HDTV

Il sistema DCT ibrido permette di realizzare co-decodificatori sia per segnali conformi alla Racc. 601, sia per segnali HDTV. Attualmente lo sviluppo di tali apparati è finalizzato alla trasmissione da punto a punto, per usi professionali, ma l'aumento delle densità di integrazione e delle velocità di elaborazione renderà possibile la realizzazione di codec di dimensioni e consumi molto ridotti.

Nel caso in cui non si pervenga ad una norma unica mondiale, l'adozione di una norma HDTV che preveda almeno un formato comune della porzione attiva dell'immagine faciliterebbe la progettazione e l'integrazione degli apparati.

Grazie alla retroazione esercitata in base al riempimento del buffer di trasmissione, il sistema si adatta automaticamente a differenti velocità di trama video ed a differenti velocità di trasmissione sul canale.

Queste considerazioni sono importanti nella definizione del formato HDTV.

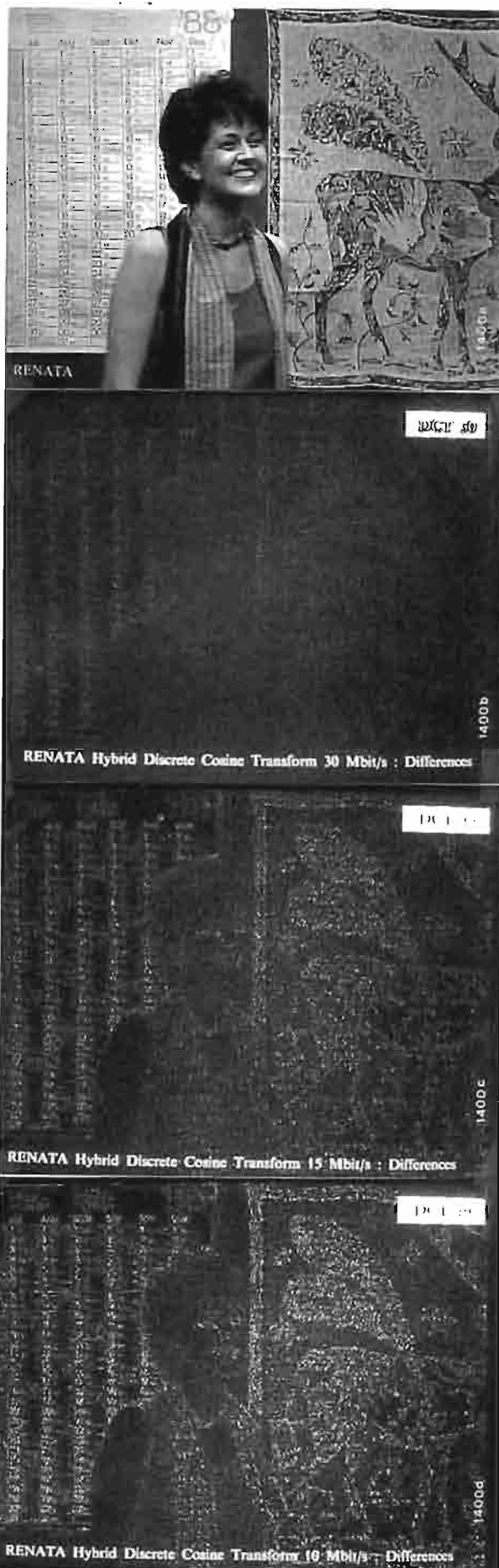



Fig. 8 — Un semiquadro tratto dalla sequenza «Renata» (ripresa presso il Centro Ricerche RAI). La prima immagine è l'originale, seguono le differenze fra essa ed i segnali co-decodificati mediante sistema DCT ibrido ai seguenti bit-rate: 30, 15, 10 Mbit/s. Viene messo in evidenza l'aumento del rumore di quantizzazione al diminuire del bit-rate e la localizzazione di tale rumore. Occorre notare che in genere è maggiormente visibile il rumore presente su aree uniformi, e meno fastidioso quello in prossimità dei contorni.

Dall'estrapolazione dei risultati ottenuti nel caso di segnali conformi alla Racc. 601 e in base ai primi risultati di simulazione su sequenze HDTV, si può ritenere che una qualità contributo sia raggiungibile, per segnali HDTV, con un bit-rate inferiore a 140 Mbit/s. Una qualità adatta alla distribuzione è ottenibile con bit-rate inferiori a 70 Mbit/s.

5.5 Altre applicazioni

I sistemi di riduzione della ridondanza per immagini hanno altri campi di applicazione, oltre alla trasmissione.

Nel caso di sistemi di videoregistrazione può occorrere una riduzione del bit-rate da memorizzare, allo scopo di limitare la velocità di scorrimento ed il consumo di nastro. Queste considerazioni valgono soprattutto nel caso di segnali HDTV.

Il sistema D1, per la registrazione di segnali conformi alla Racc. 601, non fa uso di tecniche di riduzione della ridondanza, per garantire la piena qualità, indipendentemente dal contenuto delle immagini.

Nel caso di registratori portatili, ad esempio per riprese a carattere giornalistico, è però auspicabile una riduzione degli ingombri e consumi, anche se a spese di un limitato degradamento della qualità.

Esigenze ancora più ridotte, per quanto riguarda la qualità, caratterizzano i sistemi per uso domestico. Si stanno quindi studiando sistemi basati su DCT anche per tali applicazioni, dove occorre tenere in conto le particolari esigenze legate alla tipologia degli errori introdotti dal canale e quelle imposte dalla necessità di montaggio elettronico (bibl. 10).

Studi sono in corso per ottenere riduzioni di ridondanza ancora maggiori, anche se a spese di una aumentata complessità dell'algoritmo e di una riduzione della qualità, per la memorizzazione di immagini su disco ottico. Questi troverebbero applicazioni per banche di immagini ad accesso casuale e per sistemi multimedia, eventualmente interattivi, ovvero video, audio, dati, programmi per calcolatore registrati sullo stesso supporto.

6. Conclusioni

A partire dal 1986 è iniziato presso il Centro Ricerche RAI lo studio della trasformata coseno discreta (DCT) per lo sviluppo di sistemi di riduzione della ridondanza del segnale video. L'ottimizzazione degli algoritmi è basato su simulazioni mediante calcolatore. Ciò consente di analizzare in modo dettagliato diverse possibili alternative e di combinare molteplici tecniche. Infatti è necessario sfruttare principi diversi, ciascuno dei quali consente riduzioni della ridondanza limitate oppure solo in alcune condizioni, per ottenere riduzioni fino ad 1/16 del flusso di informazione originario.

La maggior parte degli algoritmi e delle tecniche studiate erano già note da tempo, ma solo negli ultimi anni sono diventati disponibili apparati atti a permettere simulazioni che riproducono in modo fedele le situazioni reali. In appendice vengono descritte le risorse disponibili presso il Centro Ricerche per questa attività.

Gli algoritmi proposti richiedono un grande numero di operazioni a velocità elevata poiché i dati video sono originati a oltre 20 Mbyte/s nel caso di video convenzionale e almeno quattro volte più velocemente nel caso di HDTV.

La possibilità di sviluppare circuiti integrati molto complessi e veloci ha finalmente consentito di realizzare praticamente i sistemi basati su trasformata.

I risultati degli studi vengono presentati ai gruppi di lavoro internazionale interessati allo sviluppo ed alla applicazione di questi sistemi e la stretta collaborazione con l'industria, in particolare la Telettra, permette la pratica realizzazione degli apparati; alcune soluzioni originali sono oggetto di brevetti in comune.

I risultati ottenuti sono notevoli, i campi di applicazione molteplici, e mentre i primi prototipi degli apparati utilizzanti il sistema DCT ibrido, sia per la TV convenzionale che per la HDTV, sono in via di completamento, si continua a studiare ed analizzare nuove tecniche che permetteranno risultati migliori, seppure a spese di una complessità maggiore.

Ringraziamenti

Gli argomenti trattati nel presente articolo sono relativi a studi e realizzazioni effettuate in collaborazione con la Telettra S.p.A. Si ringraziano in particolare l'ing. Silvio Cucchi e l'ing. Francesco Molo.

Il lavoro svolto presso il Centro Ricerche RAI è frutto dell'attività di gruppo e quindi ringraziamo gli amici e colleghi con i quali va diviso il merito dei lavori presentati in questo articolo.

APPENDICE

Sistemi di simulazione

Più volte, nel corso dell'articolo è stata puntualizzata l'importanza delle simulazioni degli algoritmi studiati mediante elaboratori elettronici. Si ritiene quindi interessante presentare una breve descrizione dei sistemi di acquisizione, elaborazione e visualizzazione del segnale video impiegati per effettuare gli studi.

Le simulazioni devono essere effettuate necessariamente su sequenze, non su immagini fisse. Infatti alcuni difetti introdotti dall'uso delle tecniche di riduzione della ridondanza o, più in generale, di elaborazione del segnale video, sono apprezzabili solo su immagini in movimento.

In genere tutti gli organismi di ricerca utilizzano sequenze prodotte appositamente e distribuite senza fine di lucro per questo scopo. Il Centro Ricerche RAI ha prodotto alcune sequenze in formato 601 fra le più diffuse ed usate. Sono state recentemente realizzate altre sequenze in formato HDTV, che avranno, si pensa, una diffusione ancora maggiore poiché poche organizzazioni sono dotate di apparati di ripresa e registrazione adatti a questo formato. Le sequenze vengono distribuite mediante nastri magnetici secondo un formato adatto agli elaboratori elettronici.

LA RIPRESA

Poiché si prevede di utilizzare gli algoritmi studiati per la trasmissione di segnali video con qualità contributo occorre disporre di un dispositivo di ripresa, la telecamera, di alta qualità.

Uno studio TV sperimentale è stato allestito: esso è dotato di una telecamera a definizione normale, con segnale di uscita a componenti (RGB) oppure composito (PAL), e di una telecamera HDTV, multistandard con segnale d'uscita a componenti, interlacciato (1250 linee per quadro, 50 Hz e 1125 linee/60 Hz) e progressivo (625 righe, 50 Hz).

Per quanto riguarda gli apparati HDTV, sia le telecamere che i dispositivi di visualizzazione, si è ancora ad uno stadio iniziale di sviluppo; ciò implica che, malgrado siano prodotti ad altissimo contenuto tecnologico e richiedano investimenti ingentissimi, le prestazioni fornite siano inferiori a quelle che saranno ottenibili a regime.

L'ACQUISIZIONE

Le sequenze riprese dalle telecamere devono essere memorizzate sotto forma numerica.

Per quanto riguarda i segnali a definizione normale la dotazione comprende un sistema Quantel denominato Harry, basato su dischi di tipo Winchester ad accesso veloce, ed un videoregistratore D1 numerico Sony a cassette. Entrambi gli apparati sono conformi alla Racc. 601 e sono progettati per la produzione di studio.

Il sistema basato su dischi consente la registrazione di sequenze fino ad un totale di 90 secondi (è dotato di 4 dischi per un totale di circa 2 GByte e la velocità di trasferimento è superiore ai 20 MByte/s). Il segnale video può essere acquisito e montato, ovvero possono essere effettuate manipolazioni atte a riorganizzare le sequenze, quadro per quadro, e a modificarle, ad esempio mediante chroma-key. Un'unità grafica altrettanto sofisticata è associata al sistema, aumentandone la versatilità.

Il videoregistratore numerico (DVTR, Digital Video Tape Recorder) può registrare e riprodurre fino a 76 minuti di programma per cassetta, sia a 625 linee/50 Hz sia a 525 linee/60 Hz. È connesso mediante interfaccia conforme alla Racc. 656 al sistema Harry.

Nel caso di HDTV si utilizza un sistema VTE basato su memorie RAM (DVSR, Digital Video Silicon Recorder) e sviluppato appositamente per scopi di ricerca. Infatti la capacità di memoria, malgrado le dimensioni dell'apparato, è pari a 750 MByte, sufficiente per meno di 10 secondi di segnale HDTV. Il sistema è collegato direttamente ad un minicalcolatore ed è possibile l'acquisizione di immagini aventi formati diversi, sia interlacciati che progressivi.

L'ELABORAZIONE

Le sequenze numeriche vengono trasferite, ovviamente a velocità più basse rispetto a quelle di acquisizione, ai sistemi di elaborazione.

Questi consistono in tre minicalcolatori Digital del tipo μVax II. Le velocità di elaborazione di tali unità (inferiore a 1 Mips, un milione di istruzioni al secondo) non sono sufficienti ad elaborare in tempi ragionevoli la mole di dati relativi ad una sequenza video (superiori a 80 MByte) pertanto ad esse sono collegati due «Array Processor» CSPI di tipo MiniMap. Questi processori sono appositamente progettati per effettuare calcoli su dati organizzati in matrici, ed un quadro video è appunto una grande matrice di dati, e raggiungono velocità dell'ordine delle decine di MFlops (milioni di operazioni in virgola mobile al secondo).

Per valutare le prestazioni di un particolare algoritmo può essere sufficiente elaborare poche decine di quadri video, possibilmente di sequenze diverse. Ma, a scopo di effettuare dimostrazioni e, soprattutto, per effettuare prove soggettive, le uniche valide per avere una valutazione finale, occorre elaborare diverse sequenze di almeno 4 s ciascuna.

I sistemi di elaborazione sono quindi sfruttati a tempo pieno e, passando allo studio del segnale HDTV, si è evidenziata la necessità di potenziarli ulteriormente.

LA VISUALIZZAZIONE

I risultati delle simulazioni sono trasferiti nuovamente alle unità di registrazione (Harry e DVTR nel caso di segnali a definizione normale e DVSR nel caso di HDTV) e visualizzati su monitor.

Per i segnali a definizione normale sono disponibili monitor di elevatissima qualità, in grado di visualizzare anche i minimi degradamenti introdotti sulle immagini.

Nel caso di HDTV si dispone di monitor da 30 pollici, di un sistema a retroproiezione da 54 pollici e di un proiettore con schermo massimo pari a m 3,2 × m 2,4.

(3827)

BIBLIOGRAFIA

1. - BARBIERI G.: *Codifica di segnali televisivi ad alta definizione per la trasmissione numerica via satellite.* «Elettronica e Telecomunicazioni», n. 1, 1988, pagg. 29-33.
2. - DEL PERO R.: *Codifica del segnale televisivo numerico: metodi di riduzione della ridondanza.* «Elettronica e Telecomunicazioni», in questo numero.
3. - BARBIERI G.: *Codifica numerica del segnale video. Standard per gli studi televisivi.* «Elettronica e Telecomunicazioni», *n. 2, 1982, pagg. 42-52.*
4. - BARBIERI G.: *Codifica numerica del segnale video. Interfaccia per apparati di produzione televisiva.* «Elettronica e Telecomunicazioni», n. 6, 1983, pagg. 265-270.
5. - ARDITO M.: *Codifica numerica del segnale televisivo: sviluppi nella normalizzazione del videoregistratore numerico da studio.* «Elettronica e Telecomunicazioni», n. 1, 1985, pagg. 31-36.
6. - GONZALEZ R. C., WINTZ P.: *Digital Image Processing.* «Addison-Wesley Publishing Company», 1987, pagg. 267-268.
7. - *DTV-34 DCT Codec TV per sistemi di trasmissione digitale a 34 Mbit/s.* Telettra - Divisione Trasmissione, Feb. 1989.
8. - PETERSON W. W., WELDON E. J.: *Error correcting codes,* «The MIT Press», 2nd ed., 1971.
9. - BARBERO M.: *Il sistema C-MAC per la radiodiffusione diretta da satellite: codifica e multiplazione dei segnali sonori.* «Elettronica e Telecomunicazioni», n. 3, 1984, pagg. 145-148.
10. - ARDITO M., BARBERO M.: *Tecniche di montaggio dei programmi televisivi registrati su nastro magnetico.* «Elettronica e Telecomunicazioni», n. 2, 1980, pagg. 77-81.

CODIFICA DEL SEGNALE TELEVISIVO NUMERICO: ARCHITETTURA DI UN CO-DECODIFICATORE HDTV UTILIZZANTE LA DCT

M. BARBERO, S. CUCCHI(*)

SOMMMARIO — I metodi di riduzione della ridondanza per il segnale video basati sulla Trasformata Coseno Discreta (DCT) e la codifica con parole a lunghezza variabile sono stati studiati e proposti per il segnale televisivo a definizione normale. Gli stessi metodi possono essere estesi alla televisione ad alta definizione (HDTV). Un codec è in fase avanzata di sviluppo in Italia e Spagna. Nell'articolo è descritta l'architettura di tale codec caratterizzata da una struttura modulare e flessibile, in grado di adattarsi a diversi formati e a differenti velocità di trasmissione. Il codec HDTV è costituito da più moduli di elaborazione, ciascuno dei quali è in grado di codificare un segnale a definizione normale. Esso è previsto per la trasmissione punto-punto di un segnale HDTV di qualità contributo con un bit-rate di circa 140 Mbit/s, ma sarà in grado di operare, con qualità distribuzione, anche a bit-rate inferiori. È disponibile la versione inglese dell'articolo.

SUMMARY — ***Digital TV signal coding: Architecture of HDTV codec using the DCT:*** *The methods for the bit-rate reduction of video signals based on the Discrete Cosine Transform (DCT) and the Variable Length enCoding (VLC) are studied and proposed for the standard definition signal. The same methods can be applied to HDTV signals and a codec is being developed in Italy and Spain. In this article the codec architecture is described, it is characterized by a modular and flexible structure capable of adapting to different video formats and various trasmission rates. The codec is built with a few processing modules, each of them is capable of processing a standard definition signal. It is designed for point-to-point transmission of a contribution quality HDTV signal at a bit-rate of about 140 Mbit/s, but it will be capable of operation at lower rates with a distribution quality. The English version of this article is available.*

1. Introduzione

In una serie di precedenti articoli (bibl. 1, 2, 3, 4) sono state illustrate le caratteristiche di un sistema per la trasmissione del segnale video numerico utilizzante la Trasformata Coseno Discreta (DCT, Discrete Cosine Transform) e la codifica con parole a lunghezza variabile (VLC, Variable Length Coding) come metodi di riduzione della ridondanza.

In base ai risultati di tali studi sono in corso di realizzazione dei codecodificatori sia per i segnali 4:2:2 (ovvero segnali televisivi a definizione normale conformi alla Raccomandazione 601 del CCIR), sia per i segnali ad alta definizione (HDTV, High Definition TeleVision).

Gli studi degli algoritmi, le simulazioni mediante calcolatore, l'attività di coordinamento con i gruppi di lavoro e di normalizzazione internazionale che operano in questo specifico campo, la progettazione e la realizzazione circuitale sono stati e sono tuttora condotti in stretta collaborazione dal Centro Ricerche RAI, dalla Telettra S.p.A. e dagli altri enti che partecipano al progetto europeo EUREKA EU 256; tale progetto ha come finalità lo sviluppo di una famiglia di co-decodificatori per la trasmissione numerica punto-punto del segnale TV e HDTV.

Gli enti televisivi hanno l'esigenza di dotarsi di un sistema efficiente e robusto, in grado di fornire una qualità prossima a quella di studio e di funzionare anche con canali affetti da elevate probabilità d'errore, mentre le industrie hanno vincoli dovuti alla complessità e fattibilità circuitale, essendo gli apparati da progettare ai limiti delle possibilità tecnologiche attuali.

Il risultato finale è quindi un compromesso fra il meglio pensabile, o simulabile mediante calcolatore, ed il realizzabile in pratica. Gli apparati in via di completamento (dovrebbero essere disponibili i prototipi per le prove nella primavera dell'anno 1990) non saranno inizialmente dotati di tutte le opzioni che potranno, in futuro, migliorarne le caratteristiche, ma offriranno comunque prestazioni adeguate alle attuali esigenze degli enti televisivi e di gestione delle reti di trasmissione e forniranno utilissime indicazioni sui filoni di ricerca e sviluppo futuri.

2. Gli obbiettivi

Il segnale televisivo numerico 4:2:2 richiederebbe per la sua trasmissione un flusso binario (bit-rate) superiore a 160 Mbit/s, corrispondente a 16 bit/pel, cioè per elemento di immagine attiva. Le capacità disponibili sulle at-

(*) Ing. Marzio Barbero del Centro Ricerche RAI - Torino; ing. Silvio Cucchi della Telettra S.p.A. - Vimercate (MI). Dattiloscritto pervenuto alla redazione il 20 luglio 1989.

tuali reti numeriche, o comunque i costi proporzionali alle capacità impegnate, impongono che tale bit-rate venga ridotto considerevolmente. Si è dimostrato che una *qualità contributo*, ovvero sufficiente a permettere alcune operazioni di post-produzione sul segnale video, è ottenibile con un bit-rate dell'ordine di 30 Mbit/s, ovvero con circa 3 bit/pel. Una *qualità per distribuzione secondaria*, sufficiente cioè per la distribuzione del segnale all'utenza domestica, è ottenibile con circa 12 ÷ 15 Mbit/s, ovvero circa 1,2 ÷ 1,5 bit/pel (bibl. 2).

I sistemi di riduzione della ridondanza che consentono queste prestazioni sono oggetto di normalizzazione nel gruppo di lavoro CMTT/2 (bibl. 5).

Per quanto riguarda la HDTV, l'attività di normalizzazione per un sistema di trasmissione numerica è rallentata dalla mancanza di una norma per il formato di studio. Ciononostante è comunemente accettato che le stesse tecniche di riduzione della ridondanza valide per il segnale a definizione normale possano essere applicate con successo al segnale HDTV.

L'esperienza acquisita nella definizione e nella progettazione di sistemi adatti ai segnali 4:2:2 consente di prevedere che, anche nel caso della HDTV, sia possibile progettare apparati in grado di operare su segnali di formato diverso, come numero di righe per quadro e numero di quadri al secondo, e con velocità di trasmissione variabili in un'ampia gamma. Si può quindi confidare nel fatto che i codec attualmente in fase di realizzazione siano adatti, non solo ai formati HDTV per i quali sono attualmente disponibili dispositivi di ripresa e visualizzazione, ma anche ad essere impiegati, previo limitate modifiche per i formati o possibilmente l'unico formato, che verranno normalizzati.

Nella definizione dei parametri e nella progettazione si è cercato di ottimizzare il sistema in modo che sia utilizzabile sia per segnali 4:2:2 che HDTV, per i diversi formati e velocità di trasmissione, realizzando quindi un'architettura flessibile e modulare.

3. La situazione attuale

3.1 Gli apparati attualmente esistenti

Sono stati proposti per la normalizzazione e sviluppati, sotto forma di prototipi, sistemi di riduzione della ridondanza per i segnali 4:2:2 basati su tecniche predittive (DPCM, Differential Pulse Code Modulation) e su trasformate (DCT).

Il gruppo 11/7 del CCIR ha effettuato un elevato numero di prove su alcuni codec sottoposti al CMTT/2 per essere normalizzati (bibl. 6). Al momento delle prove erano disponibili i prototipi di un codec DPCM costruito in Giappone, di un codec DCT ibrido costruito in Francia e di un codec DCT puramente intra-field, prodotto di serie sviluppato in Italia (bibl. 7).

I risultati di tali prove sono brevemente riassunti in Appendice e da essi è possibile trarre alcune conclusioni. Sia la tecnica DPCM, sia quella DCT ibrida garantiscono la qualità richiesta a 45 Mbit/s. Il sistema DCT ibrido è stato provato a 34 Mbit/s con risultati generalmente buoni.

Il codec DCT puramente intra-field è stato sottoposto a prove per fornire un termine di confronto, in quanto rappresenta un prodotto industrializzato e commercializzato. Non è stato proposto come candidato alla normalizzazione perché, almeno a bit-rate pari o inferiori a 34 Mbit/s, non garantisce prestazioni sufficienti per la qualità contributo. Queste sue caratteristiche sono state confermate dalle prove: la qualità offerta è buona, ma non sufficiente nel caso in cui il segnale co-decodificato sia sottoposto a post-produzione, in particolare intarsio o chroma-key.

Per quanto riguarda il comportamento dei codec in presenza di errori e di perdita di sincronismo i risultati sono stati invece piuttosto deludenti, sia per il sistema DPCM che per quello DCT ibrido.

Per quanto riguarda soprattutto il recupero del sincronismo in caso di interruzioni del segnale in entrata al decodificatore, il comportamento migliore è stato ottenuto invece con il codec DCT puramente intra-field. Questo risultato sembra sorprendente se si considera che il codice usato per la correzione degli errori è in grado di correggere fino a due errori in una parola di 511 bit, nel caso del sistema intra-field, ed è quindi meno potente di quello usato nei sistemi DPCM e DCT ibrido, che è in grado di correggere due errori in una parola di 255 bit.

Ma ad una più attenta analisi tale risultato è ascrivibile in parte al fatto che il sistema intra-field non sfrutta la ridondanza temporale, e di conseguenza gli effetti degli errori non si propagano nei quadri video successivi, risultando quindi meno visibili. Comunque, il merito più rilevante è forse dovuto alle caratteristiche del VLC e della trama, differenti da quelli adottati negli altri codec sottoposti a prove, nel progetto delle quali è stata devoluta una particolare attenzione proprio per garantire una maggior robustezza in presenza di errori.

Tali caratteristiche sono presenti anche nell'altra versione di sistema DCT ibrido studiato e proposto al CMTT/2, ma che, a differenza della versione che è stata sottoposta a prove, non sarà disponibile prima della primavera del 1990.

Per quanto riguarda la HDTV, sono stati sviluppati sistemi, e in alcuni casi costruiti prototipi, basati su DPCM (bibl. 8, 9). L'uso della sola tecnica predittiva è meno efficiente della tecnica ibrida; pertanto per ottenere fattori di riduzione superiori a 4 è in genere necessario ricorrere alla compensazione del movimento e al sottocampionamento, cosa che comporta un aumento della complessità circuitale richiesta.

3.2 I limiti tecnologici

I circuiti richiesti per la realizzazione degli algoritmi sono notevolmente complessi, inoltre devono operare con clock di elevata frequenza. Non è pensabile, neppure a livello di prototipi, di realizzare un codec a componenti discreti, ma è necessario sviluppare circuiti con un elevato livello di integrazione. Sono quindi necessari elevatissimi investimenti, affrontabili solo nel caso in cui i componenti risultanti possano trovare un'ampia gamma di impieghi. Diventano fondamentali gli obiettivi di comunanza delle

Tabella 1.
Prestazioni dei VLSI semicustom

Tecnologia	HCMOS 1,5 μm	HCMOS 1 μm	BIMOS 1,5 μm	ECL
clock [MHz]	50	50-70	70	800
complessità [kGates]	50	100	50	5
dissipazione di potenza	2 W	2 W	2,5 W	5 W
	per f = 30 MHz per 20 kGates			5 kGts

Nota: Tecnologie per circuiti semicustom VLSI silicio disponibili attualmente, o previste a breve termine, dai principali fornitori. La tecnologia ECL ha ottime prestazioni in termini di velocità, ma basso livello di integrazione ad elevata potenza dissipata. La tecnologia CMOS consente un buon livello di integrazione con potenze dissipate accettabili. Benché la frequenza di clock possa superare i 50 MHz, soprattutto nel caso di BIMOS, è opportuno adottare un valore più conservativo, soprattutto nel caso in cui più dispositivi siano contenuti su uno stesso circuito stampato.

architetture e della modularità costruttiva per le differenti applicazioni.

La scelta della tecnologia utilizzata è fondamentale per garantire la fattibilità e l'affidabilità dei codec (bibl. 10). In Tabella 1 sono riassunte le caratteristiche principali delle famiglie VLSI semicustom attualmente disponibili in termini di livello di integrazione, massima velocità di clock e dissipazione di potenza.

Da tali dati si può trarre la conclusione che la tecnologia HCMOS consente la produzione di circuiti adatti, in termini di complessità e potenza dissipata, all'elaborazione del segnale 4:2:2; infatti è sufficiente una frequenza di clock pari a circa 30 MHz, usando la quale è garantito un certo margine tecnologico.

Nel caso di elaborazione di segnali HDTV è comunque indispensabile ricorrere a tecniche di parallelizzazione, suddividendo l'immagine da elaborare in un opportuno numero di zone.

4. I vantaggi degli algoritmi proposti

4.1 La DCT

I campioni video sono elaborati a blocchi di 8 · 8 pel (bibl. 2). L'organizzazione in blocchi facilita la progettazione di una struttura modulare, facilmente estensibile al caso HDTV.

Per la realizzazione dei codec 4:2:2 e HDTV di tipo DCT ibrido, è stato progettato e realizzato un circuito integrato in grado di effettuare in tempo reale la DCT diretta o inversa per un segnale 4:2:2. Questo dispositivo è derivato da quello attualmente usato per il codec puramente intra-field, ma offre prestazioni migliorate per quanto riguarda velocità e precisione. Realizzato in tecnologia HCMOS, ha una complessità equivalente a 32 kGate. Esso è in grado di operare ad una frequenza di clock massima pari a 32 MHz nel caso peggiore e a 60 MHz nel caso migliore. Utilizza un algoritmo di tipo veloce ed è caratterizzato da una precisione di calcolo superiore a quella richiesta dal CCITT (Comitato Consultivo di Telegrafia e Telefonia): i campioni video possono avere una precisione di 9 ÷ 11 bit, i coefficienti DCT 12 ÷ 14 bit.

4.2 La compensazione del movimento

L'applicazione della tecnica di compensazione del movimento non è strettamente indispensabile nel caso in cui si adotti la DCT per i fattori di riduzione di ridondanza presi in considerazione, ma consente di ottenere un miglioramento qualitativo, significativo nel caso di certi tipi di immagine, a parità di bit-rate oppure una riduzione del bit-rate, a parità di qualità.

Il fatto che per la compensazione del movimento occorra ricorrere a blocchi non coposizionati, appartenenti a semiquadri precedenti, comporta alcuni problemi nel caso in cui l'immagine complessiva sia suddivisa in più zone, ciascuna elaborata separatamente.

Per risolvere tali problemi si può evitare di applicare la compensazione del movimento in quei casi in cui essa implica l'uso di blocchi appartenenti ad altre zone, adiacenti a quella considerata. L'unica contropartita è una leggera perdita in efficienza, in genere del tutto trascurabile; comunque la compensazione del movimento non è applicata e questo primo prototipo.

4.3 Il VLC

Il codice a lunghezza variabile proposto (bibl. 3) è già stato sperimentato nella realizzazione del sistema DCT puramente intra-field. Ha ottime caratteristiche per quanto riguarda il recupero del sincronismo di parola, limitando la propagazione degli errori.

Le parole di codice sono costituite da coppie di bit ed esse possono essere organizzate in uscita dal codificatore VLC come rappresentato nell'esempio di figura 1.

Le parole sono prodotte con un ritmo non costante, infatti alcune rappresentano sequenze di zeri, anziché sin-

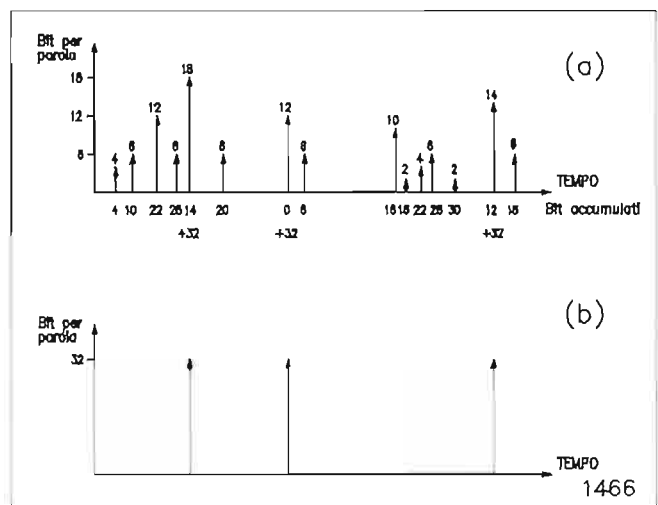



Fig. 1 — In questo esempio le parole a lunghezza variabile (a) sono generate ad un ritmo non costante, esse sono accumulate fino a quando viene generata una parola di 32 bit (b). Si osservi come le parole a lunghezza fissa siano generate con un ritmo che può essere fortemente disuniforme.

goli coefficienti, e hanno lunghezza variabile. Esse vengono accumulate in parole di lunghezza fissa, ad esempio 32 bit, in modo da poter essere trasferite, organizzate in un modo più conveniente, ai circuiti successivi. Si osservi che tali parole a lunghezza fissa sono comunque generate ad un ritmo non costante, che dipende dalla complessità relativa della porzione di immagine elaborata, e ciò dà origine a problemi di multiplazione.

In teoria la individuazione e la decodifica delle parole VLC sono le uniche operazioni che devono obbligatoriamente essere effettuate sul flusso binario seriale, alla velocità di trasmissione.

Fortunatamente la struttura a coppie di bit delle parole di codice ed il fatto che ciascun bit della coppia abbia una specifica funzione, facilita grandemente la decodifica. Il primo bit di ciascuna coppia indica se la parola di codice termina o prosegue con la coppia a cui appartiene. Di fatto, l'individuazione e la decodifica delle parole di codice può avvenire a partire dai dati organizzati, anche dal lato decodificatore, in parole di lunghezza fissa, pari a 32 bit, riducendo così la velocità operativa richiesta ai circuiti.

L'uso di particolari parole costanti per delimitare i blocchi (EOB, End Of Block) consente inoltre di individuare facilmente i dati associati ai singoli blocchi di immagine da elaborare, limitando gli effetti degli errori.

4.4 Il codice per la correzione degli errori

Per la correzione degli errori viene utilizzato un codice BCH (255,239) in grado di correggere gli errori singoli e doppi.

A tale scopo è utilizzato un circuito integrato che contiene sia il circuito di codifica che quello di decodifica per i codici BCH (511,493), BCH (255,239) e quelli accorciati da essi derivati.

Tale dispositivo è già utilizzato per il codec DCT puramente intra-field: nella versione a 34 Mbit/s è utilizzata l'opzione con parole da (511,493).

4.5 La trama a pacchetti

La struttura di trama video proposta (bibl. 4), organizzata in pacchetti di lunghezza fissa, presenta vantaggi dal punto di vista della semplicità e della robustezza agli errori, ma soprattutto consente la realizzazione modulare di codec HDTV.

Si è precedentemente sottolineata la necessità di dividere l'immagine HDTV in zone indipendenti in modo da consentire l'elaborazione parallela da parte di più unità di codifica. Grazie alla organizzazione dei dati a pacchetti, ciascun unità può operare in modo quasi totalmente indipendente dalle altre, producendo dati organizzati in pacchetti contraddistinti da un particolare indirizzo di sorgente.

L'indirizzo di sorgente contenuto nell'intestazione di pacchetto, consente quindi l'instradamento dei dati prodotti da una delle unità di codifica fino all'unità di decodifica corrispondente, preposta alla ricostruzione della zona di immagine ad essa assegnato.

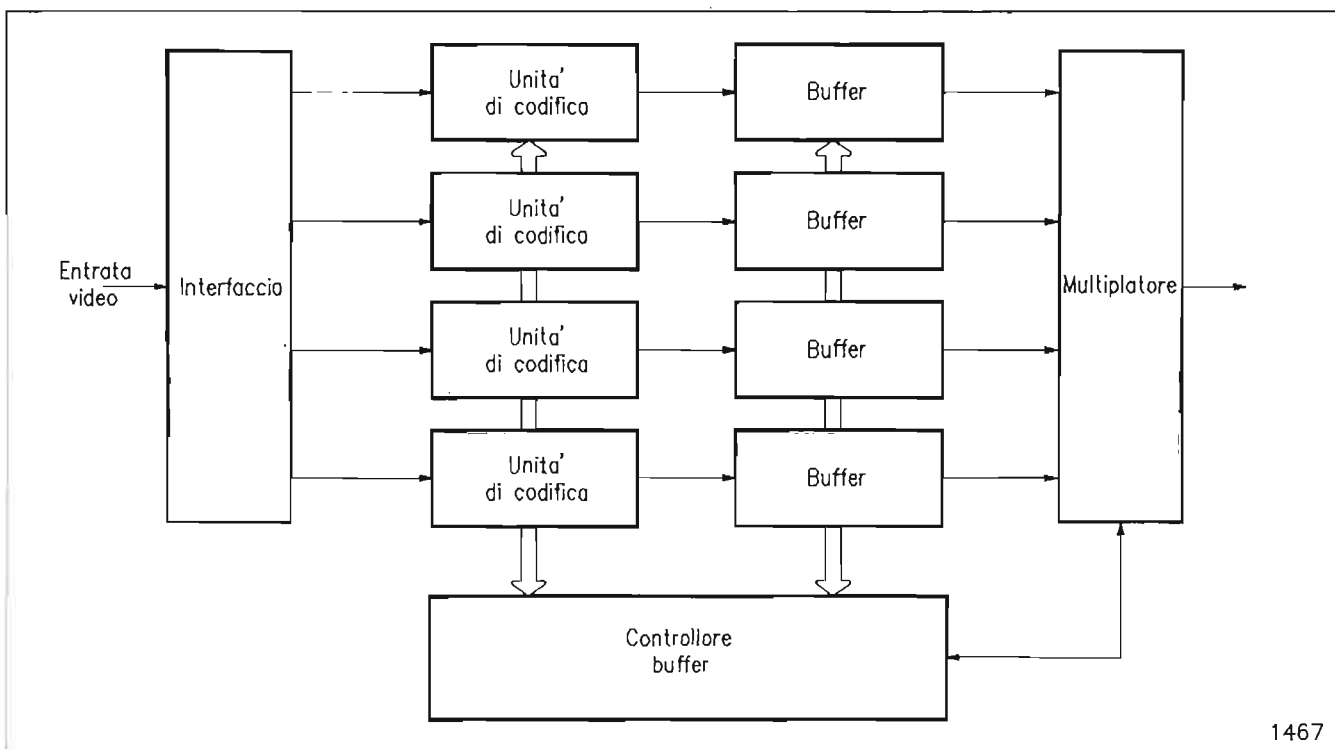



Fig. 2 — Il codificatore HDTV è composto da una interfaccia d'entrata che suddivide i campioni video attivi in N flussi, in questo esempio sono quattro, ciascuno dei quali è elaborato da un'unità di codifica. I dati risultanti sono memorizzati in un buffer e da qui prelevati, sotto forma di pacchetti, per essere multiplati e trasmessi. Il Controllore di buffer sovraintende alla organizzazione dei pacchetti, al loro prelievo ed alla determinazione del fattore di trasmissione in base all'occupazione dei buffer.

Questo processo di co-decodifica sarà più chiaro in base alla descrizione dell'architettura del codec HDTV.

5. L'architettura del codec HDTV

5.1 Lo schema a blocchi

In figura 2 è rappresentato lo schema a blocchi del codificatore HDTV.

5.2 L'interfaccia d'entrata

Il segnale in entrata viene suddiviso in N flussi, ciascuno corrispondente ad una delle N zone verticali in cui è divisa l'immagine HDTV. A ciascun flusso corrisponde una via di elaborazione.

Nel caso, ad esempio, di un formato costituito da 1440 pel per 1152 righe attive, l'immagine può essere suddivisa verticalmente in 4 zone, ciascuna costituita da un numero di pel uguale a quello presente in una immagine 4:2:2.

Quando il formato prevede un numero maggiore di pel per riga, ad esempio 1920, o un numero maggiore di righe per unità di tempo, come per i formati con scansione progressiva, il numero di flussi può essere aumentato in modo da ridurre il numero di pel attivi per ciascuna via di elaborazione.

L'interfaccia d'entrata provvede inoltre alle funzioni di recupero dei sincronismi video, generazione dei clock ed eliminazione delle porzioni non attive dell'immagine.

5.3 Le unità di codifica

I dati video appartenenti a ciascuna via vengono elaborati indipendentemente dagli altri da un'unità di codifica. Tale unità può, in linea di principio, essere basata su qualsiasi tecnica di riduzione della ridondanza. In particolare essa può essere il codificatore DCT ibrido proposto per la codifica del segnale 4:2:2, descritto in bibl. 2 ed il cui schema a blocchi è riportato in figura 3.

Le parole a lunghezza variabile in uscita da tale codificatore sono accumulate a formare parole a lunghezza fissa, ma caratterizzate da un ritmo variabile nel tempo, come descritto precedentemente.

Tali parole di lunghezza fissa potrebbero essere avviate direttamente al multiplatore di canale, ma poiché le N unità di codifica, una per ciascuna via, producono parole con un ritmo fortemente variabile e dipendente dalla complessità locale dell'immagine codificata, sarebbe praticamente impossibile individuare le singole parole in modo

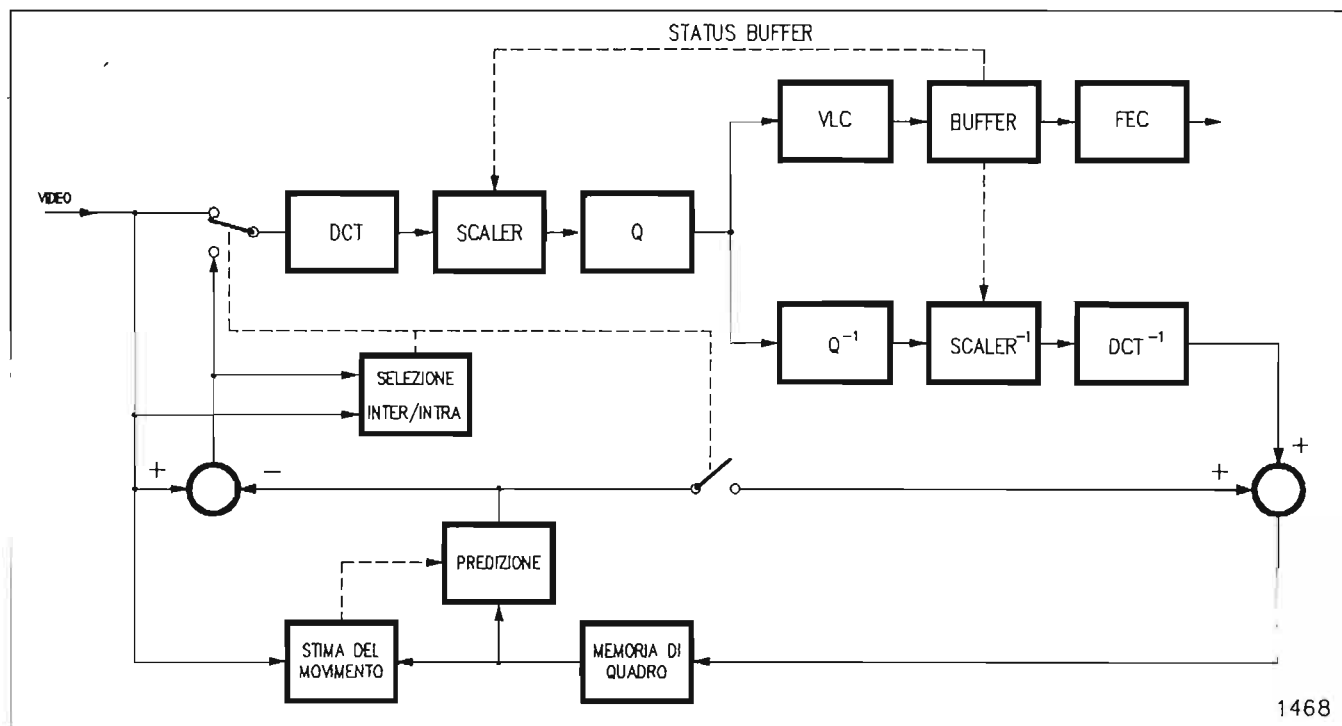



Fig. 3 — Schema a blocchi di una delle unità di codifica utilizzante come algoritmo di riduzione della ridondanza la Trasformata Coseno Discreta ibrida. I campioni in entrata, organizzati in gruppi di 8·8, vengono trasformati (DCT), moltiplicati per un fattore di scalamento (SCALER) in base al riempimento del buffer, quantizzati (Q), codificati con parole a lunghezza variabile (VLC) e memorizzati nel BUFFER dove sono organizzati in pacchetti completi di codice per la protezione dagli errori (FEC). I dati in entrata al circuito che calcola la DCT possono essere campioni video (modo intra-field) oppure differenze fra il valore di ciascun campione e un valore di predizione ottenuto a partire dai campioni appartenenti al semiquadro precedente (modo inter-field) o da quelli appartenenti al semiquadro ancora prima (modo inter-frame). In quest'ultimo caso i valori di predizione possono essere ottenuti applicando tecniche di stima e compensazione del movimento. I dati di predizione devono essere uguali, a meno di differenze dovute ad errori di trasmissione, a quelli disponibili dal lato di decodifica; per tale ragione i campioni di riferimento memorizzati nella memoria di quadro sono ottenuti mediante le operazioni inverse di quantizzazione (Q^{-1}), scalamento ($SCALER^{-1}$) e trasformata (DCT^{-1}) a partire dai dati in trasmissione.

univoco, tale da permetterne la corretta decodifica in ricezione.

Conviene quindi raggruppare le parole in pacchetti di opportune dimensioni, fornirli di una parola di intestazione che individui l'unità di codifica d'origine e, di conseguenza, quella di destinazione, e trasferire i dati al multiplatore già organizzati in pacchetti. Il formato di questi pacchetti è descritto in bibl. 4. Ciascun pacchetto contiene oltre all'intestazione e ai dati relativi al segnale video codificato, anche l'informazione relativa allo stato di riempimento del buffer di trasmissione e la ridondanza per la correzione degli errori eventualmente introdotti dal canale.

5.4 Il controllore dei buffer

Ciascuna unità di codifica è dotata di un buffer di trasmissione, che tende a riempirsi più o meno velocemente in base alla complessità della porzione di immagine codificata.

Esiste una unità di supervisione, che coordina le N unità di codifica: il controllore dei buffer. Tale unità fa si che ciascun buffer non pervenga a condizioni di svuotamento completo (underflow) o sovraccarico (overflow) e che i pacchetti in essi contenuti vengano prelevati dal multiplatore secondo l'ordine temporale con cui ha avuto inizio la loro formazione.

Ogni qual volta viene iniziata la generazione di un nuovo pacchetto all'interno di uno degli N buffer, tale evento viene segnalato al controllore di buffer. Il controllore provvede quindi a fornire l'informazione di intestazione del pacchetto, contenente l'indirizzo della via di sorgente e di destinazione, e a memorizzare le stesse informazioni al suo interno. Quando il multiplatore richiede l'invio di un nuovo pacchetto, il controllore è così in grado di determinare quale è il pacchetto disponibile più vecchio, come tempo di inizio formazione, e di ordinare al buffer della via interessata di effettuare il trasferimento del suddetto pacchetto al multiplatore.

Nel caso in cui più pacchetti inizino a formarsi contemporaneamente in due o più delle N vie, l'assegnazione arbitraria dell'ordine di trasmissione a tali pacchetti non causa problemi di underflow dal lato ricezione, grazie alla presenza del buffer nel decodificatore.

In base alla situazione di occupazione dei buffer il controllore provvede inoltre a variare il fattore di scalamento dei coefficienti DCT allo scopo di adattare il flusso di informazione generato.

Il controllo dei buffer corrispondenti alle N vie può essere effettuato secondo diverse modalità, in funzione dei compromessi realizzativi attuati.

In particolare si possono controllare i buffer singolarmente. In questo caso se ad ogni buffer si assegna una capacità pari a quella massima prevista, il controllore fa sì che la somma delle occupazioni dei vari buffer non superi mai il limite massimo e che le unità di codifica operino con la stessa precisione. Questa soluzione non consente un uso ottimale della capacità complessiva dei buffer, ma permette una notevole semplificazione del controllore.

In alternativa il controllo dei singoli buffer può essere attuato come se si trattasse di un unico buffer; oppure è possibile utilizzare un unico buffer collegato alle uscite delle N unità. Queste soluzioni consentono di ottimizzare le dimensioni del buffer a spese di una maggiore complessità del controllore e possono essere vantaggiose in caso di un valore elevato di N.

5.5 Altri servizi

Altre informazioni, ad esempio audio e dati televideo, possono essere organizzate in pacchetti del medesimo formato inviate al multiplatore.

L'intestazione di pacchetto individua il tipo di servizio a cui esso si riferisce mediante un opportuno identificatore di servizio.

All'interno di pacchetti provenienti da sorgenti di tipo differenti si possono utilizzare criteri diversi di organizzazione dei dati e codici diversi per la protezione dagli errori, in modo da adattarsi alle specifiche caratteristiche della sorgente.

La capacità complessiva assegnata a queste informazioni può essere variabile, possibilità tipica offerta dai sistemi a pacchetto, oppure fissa e pari, ad esempio, a 2048 kbit/s.

5.6 Il multiplatore

Il multiplatore può essere concepito in modo tale da trattare tutti i tipi di pacchetto e, in base alla velocità di trasmissione sul canale, interagire con il controllore di buffer per adattare il flusso di dati generati a velocità di trasmissione differenti.

La priorità assegnata in trasmissione è diversa per i diversi tipi di pacchetto, essendo, ad esempio, conveniente assegnare priorità più elevata all'informazione video contenuta nei pacchetti di tipo F, relativi quindi ai dati a lunghezza fissa, e più bassa a quelli di tipo V, corrispondenti ai coefficienti DCT codificati mediante parole a lunghezza variabile.

Il multiplatore provvede, tramite opportuna interfaccia, alla trasmissione dei dati serializzati e codificati secondo il codice di linea prescelto.

5.7 Il decodificatore

Il decodificatore ha un'organizzazione duale a quella del codificatore. In particolare i pacchetti in arrivo vengono smistati correttamente in base all'indirizzo contenuto nell'intestazione, ma il controllore del buffer deve mantenere sotto controllo lo stato di riempimento dei vari buffer e far sì che esso sia complementare a quello dei buffer nel lato codifica.

6. I vantaggi della struttura proposta

Il codec proposto ha una struttura modulare estremamente semplice.

In pratica l'elaborazione del segnale HDTV è ottenuta mediante N moduli, ciascuno dei quali è un codificatore, o decodificatore, adatto ad elaborare un segnale 4:2:2.

In questo modo si ottiene, per le differenti applicazioni, una comunanza non solo a livello di componenti, ma addirittura a livello di modulo. Ciò consente, tra l'altro, di poter effettuare le prove ed il collaudo dei codec HDTV utilizzando le stesse procedure e gli stessi segnali di prova previsti per i codec 4:2:2.

Il codec è adattabile a diversi formati video semplicemente modificando l'interfaccia d'entrata, e la corrispondente interfaccia d'uscita nel decodificatore, ed eventualmente variando il numero di unità di co-decodifica.

È inoltre possibile, utilizzando interfacce di entrata e uscita apposite, sfruttare un codec previsto per segnali HDTV allo scopo di codificare N segnali 4:2:2, assegnando dinamicamente la capacità complessiva del canale alle differenti sorgenti, in base alla complessità delle immagini elaborate oppure assegnando una capacità parziale fissa ad ognuna di esse.

Inoltre utilizzando multiplatori differenti, è possibile adattare il sistema a canali con caratteristiche e capacità diverse.

Infine può essere assegnata una capacità prefissata oppure variabile ai dati relativi ad altri servizi, grazie alla organizzazione a pacchetti.

7. Conclusioni

Si è tentato di progettare congiuntamente il sistema di riduzione della ridondanza, la struttura di trama dei dati risultanti e l'architettura dei codec per segnali video a definizione normale e ad alta definizione. In questo modo si è ottenuta un'elevata efficienza dell'algoritmo di codifica, una robustezza nei confronti degli errori introdotti dal canale, una flessibilità di impiego, pur mantenendo la complessità realizzativa e la tecnologia richiesta entro limiti tali da garantire la fattibilità in tempi e costi accettabili.

I primi prototipi saranno disponibili fra breve e consentiranno di verificare sul campo le caratteristiche previste e ottimizzate mediante simulazioni con calcolatore. Tali prototipi, comunque, saranno ad un livello di definizione tale da consentire la produzione effettiva in tempi brevi.

Ulteriori miglioramenti, già previsti e compatibili con l'organizzazione attuale verranno introdotti in un secondo tempo. È probabile che, in base all'esperienza acquisita con l'uso di questa prima generazione di codec HDTV ed in base all'attività di normalizzazione internazionale, si passi alla realizzazione di generazioni successive, possibilmente ancora più efficienti e basate su tecnologie a più elevato livello di integrazione. Si ritiene che le soluzioni proposte, anche dal punto di vista della struttura di trama, abbiano elevate potenzialità per consentire un'evoluzione futura, ma possibilmente compatibile, del sistema ed una sua applicazione per usi diversi.

Tabella 2.
QUALITÀ BASE A 34 MBIT/S

DIST.	CODEC	SEQUENZA						
		A	B	C	D	E	F `	m
4H	INTRA	43	1	5	5	2	18	12
	H-DCT	41	1	1	19	7	3	12
	RIF	4	3	5	5	5	6	5
6H	INTRA	36	1	3	1	0	12	9
	H-DCT	30	2	0	13	2	2	8
	RIF	4	4	5	4	5	5	4

Note:
(1) Il metodo di votazione è quello del doppio stimolo con scala di qualità continua, successivamente suddivisa in 100 divisioni, le cifre riportate si riferiscono a centesimi. La distanza di visione è pari 4 volte (4H) oppure 6 volte (6H) l'altezza dello schermo.
(2) Numero degli osservatori: 150.
(3) I codec confrontati sono: INTRA (codec DCT intra-field a 34 Mbit/s, Italia), H-DCT (DCT ibrida a 34 Mbit/s, Francia). RIF indica la sequenza di riferimento.
(4) Le sequenze di prova sono: Diva with Noise (A), Tempete (B), Flower Garden (C), Kiel Harbour (D), Rotating Disc (E), Mobile and Calendar (F).
(5) La media calcolata su tutte le sequenze è indicata come m.

Tabella 3.
QUALITÀ BASE A 45 MBIT/S

DIST.	CODEC	SEQUENZA						
		G	B	H	D	l	F	m
4H	DPCM	1	1	4	0	4	1	2
	H-DCT	0	−1	−1	0	2	1	0
	RIF	3	3	4	4	4	5	4
6H	DPCM	0	0	3	0	1	3	1
	H-DCT	1	−1	2	0	0	−1	0
	RIF	6	4	6	5	5	6	5

Note:
(1) vedere nota (1) di Tabella 2.
(2) Numero degli osservatori: 72.
(3) I codec confrontati sono: DPCM (codec DPCM a 45 Mbit/s, Giappone), H-DCT (DCT ibrida a 45 Mbit/s, Francia). RIF indica la sequenza di riferimento.
(4) Le sequenze di prova sono: Ciao (G), Tampete (B), Popple (H), Kiel Harbour (D), Balls of wool (l), Mobile and Calendar (F).
(5) Vedere nota (5) di Tabella 2.

APPENDICE

Prove sui codec 4:2:2

Nella primavera 1989 sono state condotte prove sui codec sottoposti al CMTT/2 al fine della definizione della norma di trasmissione del segnale 4:2:2 con qualità contributo. Le prove sono state effettuate a cura del Gruppo AD-HOC BCT (Bit-rate reduction Codec Testing) mediante alcune delle sequenze raccolte e distribuite dal Gruppo AD-HOC TPS (Test Picture Sequences); entrambi i gruppi dipendono dal IWP 11/7 del CCIR e la metodologia delle prove è stata definita in collaborazione con lo IWP 11/4.

Sono stati sottoposti a prove nelle località indicate i seguenti codec: DPCM a 45 Mbit/s giapponese (Tokyo), DCT ibrida a 45 Mbit/s e a 34 Mbit/s francese (Rennes), DCT intra-field a 34 Mbit/s italiano (Monaco). I codec a 45 Mbit/s operavano su segnali a 525 righe/60 Hz, quelli a 34 Mbit/s su 625 righe/50 Hz. I codec DPCM e DCT ibrido sono prototipi, mentre il codec DCT intra-field è prodotto dalla Telettra S.p.A. ed è disponibile in commercio.

Le prove tendevano ad accertare: la qualità base, la

qualità dopo operazioni di post-produzione (intarsio o chromakey e slow-motion), la protezione verso gli errori.

Come sorgente delle sequenze utilizzate per le prove si è usato un videoregistratore numerico 4:2:2 (formato D1). Le sequenze risultato delle prove sono state organizzate in nastri D1 secondo le procedure per prove soggettive e distribuite a 7 laboratori nel caso del formato 625 righe/50 Hz e a 3 laboratori per quello 525 righe/60 Hz; il Centro Ricerche RAI ha effettuato le prove per entrambi i formati.

Le tabelle e la figura che seguono riassumono i risultati delle prove.

I risultati delle prove soggettive di qualità base e dopo post-produzione indicano la differenza in qualità rispetto a quella di riferimento, cioè dell'originale, pertanto i valori bassi indicano degradamento minimo e il degradamento massimo è rappresentato dal valore 100. Si richiede (bibl. 11) che la differenza in qualità per qualsiasi delle sequenze di prova sia non superiore a 12 per la qualità base e 18 per la qualità dopo post-produzione.

Dall'esame dei risultati di Tabella 2 emergono le considerazioni che seguono.

La sequenza «Diva with noise» (A) è particolarmente critica, in pratica è stata proposta allo scopo di evidenziare i difetti dei sistemi di riduzione della ridondanza (vedere bibl. 2, paragrafo 3.6).

La sequenza «Mobile and calendar» (F) è caratterizzata da dettaglio e molta informazione di crominanza, è quindi di difficile codifica da parte del sistema DCT intra-field, mentre il sistema ibrido offre buone prestazioni. È inoltre noto che tale sequenza trae un forte vantaggio dall'uso della compensazione del movimento.

Nel caso della sequenza «Kiel Harbour» (D) il codec ibrido presentò durante le prove dei problemi di rigenerazione del clock. Il risultato indicato deve quindi essere ignorato.

Entrambi i codec presentano buone prestazioni nel caso delle restanti sequenze.

Dai risultati di Tabella 3 si trae la conclusione che la qualità offerta da entrambi i codec, DPCM e DCT ibrido, è buona a 45 Mbit/s. I codec non sono stati provati con la sequenza più critica, ovvero «Diva with noise» (A).

Dalla Tabella 4 si desume una prestazione molto migliore del sistema DCT ibrido rispetto a quello DCT intra-field nel caso di intarsio o chroma-key. Ciò è particolarmente evidente nel caso della sequenza «Master» (L).

A 45 Mbit/s, come si desume dalla Tabella 5, entrambi i sistemi DCT ibrido e DPCM offrono buone prestazioni. Anche in questo caso non era disponibile la sequenza risultata più critica nel caso del 34 Mbit/s, cioè «Master» (L).

La figura 4 mostra una buona prestazione dei codec DPCM e DCT ibrido con probabilità d'errore inferiori a 10^{-5}, ma un rapido peggioramento al crescere della stes-

Tabella 4.
POST-PRODUZIONE A 34 MBIT/S

DIST.	CODEC	CHROMA-KEY SEQUENZA I	L	G	mc	SLOW-MOTION SEQUENZA E	D	M	ms
4H	INTRA	3	28	0	11	−2	6	0	1
	H-DCT	4	10	2	5	2	3	−3	1
	RIF	4	5	4	4	3	4	3	3
6H	INTRA	1	24	1	9	1	2	0	1
	H-DCT	4	7	1	4	1	1	−5	−1
	RIF	5	4	4	4	4	3	2	3

Note:
(1) vedere nota (1) di Tabella 2.
(2) Numero degli osservatori: 126.
(3) vedere nota (3) di Tabella 2.
(4) Le sequenze di prova sono: Balls of wool (I), Master (L), Ciao (G), Rotating Disc (E), Kiel Harbour (D), Cruising (M).
(5) La media calcolata sulle sequenze su cui è stato effettuato l'intarsio (CHROMA-KEY) è indicata come mc, quella sulle sequenze per cui si è praticato il movimento rallentato (SLOW-MOTION) è indicato come ms.

Tabella 5.
POST-PRODUZIONE A 45 MBIT/S

DIST.	CODEC	CHROMA-KEY SEQUENZA H	B	G	mc	SLOW-MOTION SEQUENZA N	D	M	ms
4H	DPCM	1	2	−1	1	−1	0	0	0
	H-DCT	1	0	0	1	1	−3	1	0
	RIF	4	4	3	4	4	3	5	4
6H	DPCM	0	2	0	0	0	−1	0	−1
	H-DCT	0	0	1	0	1	−2	1	0
	RIF	4	6	4	5	5	5	5	5

Note:
(1) vedere nota (1) di Tabella 2.
(2) Numero degli osservatori: 69.
(3) vedere nota (3) di Tabella 2.
(4) Le sequenze di prova sono: Popple (H), Tempete (B), Ciao (G), Shinjuku (N), Kiel Harbour (D), Cruising (M).
(5) vedere nota (5) di Tabella 4.

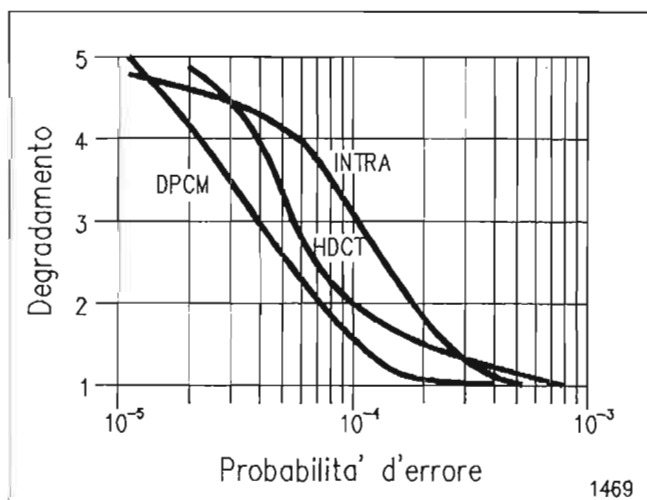



Fig. 4 — Prestazioni dei seguenti codec: DPCM (prototipo DPCM a 45 Mbit/s, Giappone); H-DCT (prototipo DCT ibrida a 34 Mbit/s, Francia); INTRA (codec DCT intra-field a 34 Mbit/s, Italia, prodotto dalla Telettra S.p.A.). I risultati sono espressi come gradi di degradamento (scala CCIR a 5 livelli; 5 = impercettibile, 4 = percettibile ma non disturbante, 3 = leggermente disturbante, 2 = disturbante, 1 = molto disturbante) in funzione della probabilità di errore sul canale. Il numero di osservatori è pari a 30 per i sistema a 34 Mbit/s e a 72 per quelli a 45 Mbit/s.

sa. Il codec DCT intra-field presenta una caduta in prestazioni più graduale.

È stato inoltre valutato il tempo di recupero del sincronismo nel caso in cui si presenti un'interruzione del segnale pari a 50 ms. I tempi di recupero per i diversi codec sono stati i seguenti:

DCT intra-field	2 semiquadri
DCT ibrido	16 semiquadri
DPCM	16 semiquadri

Solamente il codec DCT intra-field ha soddisfatto i requisiti di recupero del sincronismo prefissati.

(3846)

BIBLIOGRAFIA

1. - DEL PERO R.: *Codifica del segnale televisivo numerico: metodi di riduzione della ridondanza.* «Elettronica e Telecomunicazioni», n. 1, 1989, pagg. 9-16.

2. - BARBERO M., STROPPIANA M.: *Codifica del segnale televisivo numerico: sistemi di riduzione della ridondanza mediante l'uso della trasformata coseno discreta.* «Elettronica e Telecomunicazioni», n. 1, 1989, pagg. 17-29.

3. - BARBERO M., BELLORA R., STROPPIANA M.: *Codifica del segnale video numerico: Codici a lunghezza variabile applicati ai coefficienti DCT.* «Elettronica e Telecomunicazioni», in questo numero.

4. - BARBERO M., DEL PERO R., GIROMINI P.: *Codifica del segnale televisivo numerico: struttura di trama per un sistema basato su DCT.* «Elettronica e Telecomunicazioni», in questo numero.

5. - IWP CMTT/2: *Revised version of Draft Report AD/CMTT: Digital Transmission of Component-Coded Television Signal at 30-34 Mbit/s and 45 Mbis/s.* Giugno 1989.

6. - *Progress Report by AHG-BCT (Ad-Hoc-Group Bit-rate reduction Code Testing).* doc. IWP 11/7-249, Giugno 1989.

7. - *DTV-34 DCT Codec TV per sistemi di trasmissione digitale a 34 Mbit/s.* Telettra - Divisione trasmissione, Febbraio 1989.

8. - MATSUMOTO S., SAITO M., MURAKAMI H.: *120/140 Mbps Intrafield DPCM System for Digital Transmission of HDTV Programs.* Second International Workshop on Signal Processing of HDTV, L'Aquila 1988.

9. - SAWADA K., YASHIMA Y.: *High Efficiency Coding of HDTV Signals at the BroadBand ISDN H4 Rate.* Second International Workshop on Signal Processing of HDTV, L'Aquila 1988.

10. - BARBIERI G., MOLO F., TEJERINA J.L.: *A Modular and Flexible Video Codec Architecture for Application to TV and HDTV.* 16th International TV Symposium Montreux Switzerland, Broadcast Session, pagg. 410-420, Giugno 1989.

11. - IWP 11/7: *Draft Report AG/11 User Requirements for Digital Television Transmission.* Giugno 1989.

CODIFICA DEL SEGNALE TELEVISIVO NUMERICO: STRUTTURA DI TRAMA PER UN SISTEMA BASATO SU DCT

M. BARBERO, R. DEL PERO, P. GIROMINI*

SOMMARIO - Il presente articolo illustra l'organizzazione di trama progettata al fine della trasmissione di segnali video numerici codificati mediante l'uso della Trasformata Coseno Discreta (DCT). Si analizzano gli effetti degli errori di trasmissione sia sotto forma di degradamento qualitativo delle immagini ricostruite, sia come perdita di sincronismo dell'informazione video. Per ovviare a tali problemi possono essere utilizzati codici per la correzione degli errori: vengono valutate le prestazioni ottenibili con codici ciclici abbinati a tecniche di «interleaving». Si passa quindi alla descrizione di una struttura di trama che consente di ottenere un'alta efficienza unitamente ad un'elevata robustezza agli errori. Significativo è l'utilizzo della tecnica di mascheramento degli errori che permette di ridurre il degradamento con l'introduzione di una ridondanza limitata grazie all'uso di codici ciclici a rivelazione degli errori (CRC). Di questa tecnica vengono anche analizzati i limiti intrinseci. I risultati presentati sono basati su simulazioni effettuate presso il Centro Ricerche RAI su sequenze video mediante l'uso di calcolatore. Sono attualmente in fase avanzata di sviluppo codec utilizzanti l'organizzazione dei dati descritti, che sono oggetto di una proposta di normalizzazione presso il CMTT. È disponibile la versione inglese dell'articolo.

SUMMARY - ***Digital TV signal coding: framing structure for a system based on DCT.*** *The present article describes the frame organisation designed for the transmission of digital video signals coded by using the Discrete Cosine Transform (DCT). The effects of transmission errors, both as picture quality impairment and as synchronisation loss of video information, are examined. To solve such problems error correcting codes can be used: the performance of cyclic codes together with interleaving techniques is evaluated. The a frame structure which is characterised by a high efficiency and robustness against errors is described. Significant is the use of the error concealment technique: it allows the reduction of subjective impairment due to errors by usign cyclic codes for error detection (CRC). The intrinsic limits of such a technique are analysed. The presented results are based on computer simulations carried out at the RAI Research Centre on video sequences. Codecs using the described data organisation are at present being developed and the structure is subject of a proposal to CMTT for standardisation. The English version of this article is available.*

1. Introduzione

In un articolo precedente (bibl. 1) sono state descritte le ricerche svolte presso il Centro Ricerche RAI volte alla definizione di un codificatore del segnale televisivo basato sull'algoritmo di Trasformata Coseno Discreta (DCT, Discrete Cosine Transform) ibrida. Queste ricerche hanno portato alla definizione dello schema funzionale del codecodificatore descritto nello stesso articolo. Sorge il problema di trasmettere il segnale così codificato, di organizzare opportunamente le informazioni e proteggere le stesse dagli errori di trasmissione.

In questo articolo si affronta il problema della definizione di una struttura di trama che presenti buone caratteristiche dal punto di vista del recupero dei sincronismi, elevata robustezza in presenza di errori di trasmissione e capacità di limitare il degradamento qualitativo da essi introdotto.

Tale struttura di trama è adatta sia alla trasmissione del segnale televisivo a definizione normale, sia alla trasmissione di segnali ad alta definizione (HDTV, High Definition Television). Nel seguito si individuerà con il termine 4:2:2: il segnale video numerico a componenti conforme alla Raccomandazione 601 del CCIR; tale denominazione fa riferimento al rapporto fra le frequenze di campionamento della componente di luminanza e delle due componenti di crominanza. Nel caso di HDTV il formato è tuttora in discussione presso gli organismi di normalizzazione.

L'argomento della codifica del segnale video per la sua trasmissione è oggetto di studi nell'ambito del progetto europeo di ricerca EUREKA EU-256 a cui partecipano, oltre al Centro Ricerche RAI, l'ente televisivo spagnolo RTVE, l'Universidad Politecnica de Madrid e, in qualità di industrie la Telettra S.p.A. e la Telettra Espana S.A. I risultati di questi studi sono oggetto di contributi ad organismi internazionali per la normalizzazione, quali il gruppo di lavoro CMTT/2 (bibl. 2).

(*) Ing. Marzio Barbero, ing. Roberto Del Pero del Centro Ricerche RAI - Torino; ing. Pierangelo Giromini del Politecnico di Torino. Dattiloscritto pervenuto alla redazione il 20 luglio 1989.

2. Struttura del segnale video

2.1 Metodi per la riduzione della ridondanza

Per comprendere l'organizzazione del segnale al fine della sua trasmissione è opportuno riesaminare brevemente le operazioni che su questo vengono effettuate; per maggiori dettagli si fa comunque riferimento ad altri articoli che meglio affrontano questi argomenti e, in particolare, a bibl. 1 per le operazioni di codifica e bibl. 3 per la descrizione del codice a lunghezza variabile adottato.

La DCT viene calcolata su blocchi quadrati di 8 pel (elementi di immagine) di lato usando un algoritmo di codifica di tipo ibrido, che adotta cioè sia metodi predittivi che di trasformazione. In base a criteri di minimizzazione della quantità di dati prodotti per unità di tempo (bit-rate) vengono codificati i campioni video appartenenti a ciascun semiquadro (modo intra-field), le differenze tra i campioni appartenenti al semiquadro attuale e quelli appartenenti al semiquadro immediatamente precedente (modo inter-field) oppure le differenze rispetto al semiquadro precedente coposizionato (modo inter-frame).

L'operazione di quantizzazione converte questi valori in un numero prefissato di livelli secondo una legge che è essenzialmente uniforme.

I livelli di uscita del quantizzatore sono codificati mediante un codice a lunghezza variabile (VLC, Variable Length Code); questa operazione ha lo scopo di adattare il più possibile il numero medio di bit di codifica all'entropia della sorgente.

Il codice usato è citato in letteratura come B_2- code in quanto l'unità minima di codifica è costituita da una coppia di bit.

Il primo bit è detto bit di continuità ed indica semplicemente se la parola di codice prosegue o termina dopo la coppia a cui appartiene; in questo modo si possono identificare le singole parole di codice indipendentemente dalla loro decodifica, opportunità non possibile per altri codici a lunghezza variabile, quali ad esempio quello di Huffman. L'uso di questo codice risulta vantaggioso al fine di semplificare le operazioni di decodifica e di facilitare il recupero del sincronismo di parola.

Il secondo bit di ogni coppia è invece un bit di informazione.

Per rendere più efficiente l'operazione di codifica della sorgente, in entrata al blocco VLC vengono codificati non solo i coefficienti DCT, ma anche le sequenze di zeri consecutivi trattate come simboli distinti dell'alfabeto di sorgente. Questa scelta risulta vantaggiosa in quanto l'operazione di trasformazione, eleminando la correlazione tra i campioni, altera la loro distribuzione statistica in modo tale che essa presenti un picco centrato attorno al valore zero.

Un'opportuna scansione dei coefficienti permette di avere con alta probabilità sequenze di zeri consecutivi specie in corrispondenza dei coefficienti a frequenza più elevata. In base a questa considerazione si preferisce non inviare la parola corrispondente all'ultima sequenza di zeri, se presente, ma utilizzare un simbolo opportuno per indicare la fine del blocco, tale simbolo è indicato come EOB (End Of Block). Questa scelta permette un'ulteriore riduzione della ridondanza.

2.2 Organizzazione dei dati video

Come detto precedentemente, l'operazione di trasformazione DCT avviene su un *blocco* quadrato di 8 pel di lato. Questo rappresenta l'unità minima di organizzazione dei dati video e viene delimitato, dopo la codifica VLC, dagli EOB.

Il livello successivo di organizzazione è il *quadriblocco*: poiché le due componenti di crominanza risultano sottocampionate orizzontalmente, un blocco di codifica delle componenti di crominanza di 8 campioni di lato interessa una superficie equivalente a due blocchi di luminanza adiacenti. Il quadriblocco è pertanto costituito dai dati relativi a due blocchi di luminanza adiacenti e ai due blocchi di crominanza corrispondenti e rappresenta quindi l'informazione completa relativa ad una superficie di 16×8 pel. I dati relativi a ciascuno dei due blocchi di luminanza di un quadriblocco sono trasmessi intercalati con quelli dei due blocchi di crominanza.

Il livello di organizzazione ancora superiore è la *striscia* che rappresenta 8 righe video attive consecutive. Nel caso di segnale televisivo 4:2:2; per il quale la riga attiva video è composta da 720 campioni, una striscia risulta costituita da 45 quadriblocchi.

Infine ciascun quadro televisivo è costituito da due semiquadri, per i sistemi interlacciati, e nel caso del sistema 4:2:2 un semiquadro contiene 36 strisce (per il formato a 625 righe/50 Hz) oppure 31 strisce (525 righe/60 Hz).

A ciascun livello di organizzazione dei dati video corrispondono sia dati a lunghezza variabile, derivati dalla codifica VLC dei coefficienti DCT, sia dati a lunghezza fissa.

A livello di singolo quadriblocco vengono codificati i modi di trasmissione relativi ai due blocchi di luminanza ed alle due componenti di crominanza congiuntamente. Per ciascun quadriblocco si indica anche se viene usata la compensazione del movimento e, in caso affermativo, il vettore da utilizzare.

Il fattore di trasmissione è invece riferito ad ogni striscia e, separatamente, alle componenti di luminanza e crominanza. Questa informazione è necessaria al ricevitore per calcolare il fattore di scalamento per i coefficienti DCT.

Esistono poi un certo numero di informazioni che interessano globalmente un semiquadro, quali le componenti dei vettori movimento ed altre informazioni di tipo più generale e ripetute a livello di semiquadro.

2.3 Sincronizzazione del segnale video

Per poter ricevere e decodificare correttamente il segnale video numerico così organizzato è necessario definire vari livelli di sincronizzazione.

Il primo livello è costituito dal sincronismo di bit che deve essere recuperato dal demodulatore.

Il sincronismo di parola, relativo alla codifica dei coefficienti DCT, è ottenuto a partire dalla struttura del codice a lunghezza variabile ed in particolare del bit di continuità.

Esiste poi la necessità di recuperare un sincronismo di fine blocco, essendo variabile il numero dei coefficienti codificati, e questo viene ricavato dalla presenza degli EOB.

L'ultimo livello di sincronismo riguarda l'intero quadro video.

Il recupero dei sincronismi deve poter avvenire anche in presenza di errori di trasmissione. La perdita del sincronismo di blocco può comportare un notevole degradamento della qualità dell'immagine decodificata, degradamento che interessa tutti i quadriblocchi finché il sincronismo stesso non viene recuperato.

Appare quindi evidente l'importanza di rivelare la perdita dei sincronismi di blocco e di ridurre al minimo il tempo per il loro recupero. A questo scopo si può convenientemente sfruttare la presenza di due EOB distinti (indicati come EOB_0 ed EOB_1) e generare sequenze secondo un ordine predefinito dei due simboli.

La sequenza di ripetizione deve essere deterministica, predefinita e presentare un periodo di ripetizione elevato.

Queste esigenze richiedono che la funzione di autocorrelazione, relativa alla sequenza che genera i due simboli, sia il più possibile concentrata in modo da garantire un'elevata capacità di identificazione del simbolo all'interno di una sequenza anche lunga.

Per rispondere a queste esigenze si fa ricorso ad un generatore pseudocasuale per ricavare la successione dei due simboli; la stessa sequenza deve essere disponibile sia al codificatore che al decodificatore.

Si prevede di generare la sequenza di EOB tramite due generatori pseudocasuali. Il primo dei due è costituito da un registro a scorrimento (*shift register*) costituito da 4 celle ed ha una periodicità pari a $2^4 = 16$; la sequenza generata è la seguente:

0000 1111 0101 1001

oppure la sequenza complementare:

1111 0000 1010 0110

dove 0 corrisponde all'invio di EOB_0 e 1 a quello di EOB_1.

Vengono emesse la prima sequenza oppure la sua complementata in base ai bit forniti dal secondo generatore pseudocasuale che ha un polinomio generatore di grado 11 e periodicità pari a $2^{11} - 1$. Complessivamente la periodicità della sequenza è pari a $2^{15} - 2^4$; lo stato iniziale del secondo generatore è caricato all'inizio di ogni quadro.

In ricezione è possibile rivelare facilmente la perdita o la creazione di EOB a causa di errori di trasmissione basandosi sulla loro periodicità con fattore 16.

In base al fatto che venga ricevuta la prima sequenza oppure la sua complementata, viene caricato uno 0 oppure un 1 nel registro corrispondente al secondo generatore pseudocasuale; in tal modo esso viene sincronizzato.

La sincronizzazione avviene quindi in due passi e ciò permette di avere tempi brevi di recupero del sincronismo sulla sequenza più breve e di poter rivelare la presenza di errori sugli EOB, senza dover attendere la ricezione di un intero quadro. Al tempo stesso ogni EOB è unicamente identificato all'interno di un quadro.

Da simulazioni effettuate si è visto che con questa struttura il sincronismo video può essere mantenuto anche con tassi di errore dell'ordine di 10^{-2}.

3. Cause ed effetti degli errori

È opportuno definire una struttura di trama per il segnale video tenendo in considerazione la distribuzione statistica degli errori e gli effetti che essi producono; in questo modo è possibile proteggere adeguatamente l'informazione trasmessa, pur mantenendo al minimo la ridondanza aggiunta.

A causa del rumore presente sul canale di trasmissione possono essere generati degli errori sui bit trasmessi; se si suppone di descrivere il rumore come un processo additivo gaussiano bianco, gli errori risultano essere statisticamente indipendenti: questo è il caso in cui il canale numerico può essere rappresentato come un canale binario simmetrico. Questo modello risulta senz'altro adeguato nel caso in cui la trasmissione del segnale video avvenga su fibra ottica dove, inoltre, la probabilità di errore ha valori molto bassi. Diverso è il caso di trasmissione punto-punto su ponte radio o via satellite dove la probabilità di errore è generalmente più elevata e, a causa di anomalie trasmissive, quali affievolimenti o *fading*, o per l'uso di schemi di modulazione con codifica differenziale, si possono avere probabilità di errore molto elevate concentrate in intervalli di tempo ristretti. In questo caso la distribuzione degli errori non è più uniforme, ma si hanno burst di errore, cioè di intervalli di tempo in cui il canale presenta una probabilità di errore sul bit (BER, Bit Error Rate) pari ad 1/2.

Attualmente è specificato (bibl. 4) come requisito per la trasmissione televisiva numerica che il degradamento dell'immagine rispetto all'originale 4:2:2 sia accettabile a tassi di errore di 10^{-4} (la Raccomandazione G821 del CCITT indica 10^{-3} come limite operativo normale). Ciò può includere burst di errore fino a lunghezze di 30 bit, benché le caratteristiche dei burst di errore siano ancora da specificare.

Nella definizione degli schemi di protezione dagli errori è necessario prevedere entrambi questi fenomeni.

Analizziamo ora quali sono gli effetti che un errore produce sulle varie informazioni trasmesse.

Una prima distinzione va fatta tra le informazioni aggiuntive, tipicamente indicate come dati a lunghezza fissa, ed i coefficienti DCT, codificati con parole a lunghezza variabile.

Fra le prime vi sono ad esempio il fattore di trasmissione, le coordinate dei vettori movimento, la codifica dei modi di trasmissione. È evidente che la presenza di errori comporta un sensibile degradamento della qualità dell'immagine ricevuta la cui entità dipende anche dal tipo di parametro. Ad esempio l'utilizzo di un fattore di trasmissione diverso da quello usato in fase di codifica porta ad un diverso scalamento inverso dei coefficienti appartenenti ad una striscia. Se invece sono alterati i valori delle coordinate dei vettori movimento, si ha un'errata traslazione dei blocchi di predizione all'interno dell'immagine e quindi una ricostruzione errata della immagine stessa. Se infine sono interessati dagli errori i modi di trasmissione, si ha, anche in questo caso, l'adozione in ricezione di una predizione diversa da quella originaria.

È evidente come tutti questi errori possano risultare ampiamente visibili sull'immagine ricevuta e come quindi sia necessario adottare schemi molto efficienti di protezione dagli errori.

Se ora passiamo ad analizzare gli effetti degli errori sui coefficienti DCT è opportuno considerare la presenza di un'operazione di codifica a lunghezza variabile e che questa è effettuata sia per i singoli coefficienti, sia per le sequenze di zeri consecutivi; a seconda della funzione del bit affetto da errore si possono verificare quindi situazioni differenti fra loro.

Innanzitutto se l'errore interessa un bit di continuità, ciò può comportare la creazione di due parole a partire da una, oppure la fusione di due parole in una sola.

Si noti che nel secondo caso il decodificatore può es-

```
   COEFFICIENTI DCT:                  403                     -23
   CODIFICA VLC:           11 11 10 11 11 10 10 00     10 10 10 11 11 01

a) ERRORE:           11 11 10 01     11 10 11 10 00     10 10 10 11 11 01
   DECODIFICA:            6              0 - 21                -23

b) ERRORE:           11 11 10 11 11  10 11 10 10 10 10 10 11 11 01
   DECODIFICA:                       impossibile

c) ERRORE:           11 11 10 11 11 11 11 10 00       10 10 10 11 11 01
   DECODIFICA:                411                          -23             1461
```

Fig. 1 - Esempio di codifica di due coefficienti mediante VLC: effetto dovuto all'introduzione di un errore. Nel caso a) l'errore è su un bit di continuità e pertanto a partire da una singola parola ne vengono generate due sbagliate con conseguente errato posizionamento dei coefficienti appartenenti al blocco e codificati mediante le parole successive a quella affetta da errore. Nel caso b) l'errore sul bit di continuità comporta la fusione di due parole di codice, in questo caso particolare la parola risultante non può essere decodificata perché di lunghezza superiore alla massima, viene quindi rivelata la presenza di un errore. Nel caso c) l'errore è su un bit di informazione e comporta l'errata decodifica del coefficiente corrispondente.

sere impossibilitato ad effettuare la decodifica, se la parola risultante ha una lunghezza superiore a quella massima, pari a 18 bit e viene quindi rivelata la presenza di errori.

Se invece l'errore interessa un bit di informazione, rimane invariata la corretta divisione delle parole, ma viene decodificato un simbolo diverso da quello originario. Può quindi accadere che venga variato il valore di un coefficiente DCT oppure che un coefficiente venga decodificato come una sequenza di zeri o viceversa.

In figura 1 sono riportati alcuni esempi di errata decodifica dovuta ad un singolo errore.

In generale la presenza di errori può comportare, oltre alla variazione dei valori dei coefficienti DCT, anche la variazione del loro numero o l'errato posizionamento di coefficienti non affetti da errore. Si noti che, se il numero di coefficienti ricostruiti rimane inferiore a 64, il decodificatore effettua egualmente la decodifica ponendo a zero i coefficienti rimanenti, essendo noto che l'ultima sequenza di zeri non viene trasmessa, ma viene anticipata la trasmissione dello EOB. Se invece si supera il numero di 64 coefficienti, non è più possibile effettuare una corretta decodifica e viene rivelata la presenza di errori.

Consideriamo ora l'effetto degli errori sugli EOB. Se si perde un singolo EOB, l'errore interessa i due blocchi di codifica adiacenti ad esso che risultano non essere più distinti; si è quindi in presenza di un errore che si propaga al di fuori del blocco. Se invece si perde un numero maggiore di EOB viene a mancare la possibilità di decodificare i coefficienti in modo corretto fino a quando il sincronismo non viene recuperato. Infine può accadere che errori interni alla codifica dei coefficienti DCT portino alla generazione di falsi EOB; anche questa situazione può portare alla perdita del sincronismo.

Relativamente agli errori sui coefficienti DCT, è necessario considerare che questi possono propagarsi nei semiquadri successivi se il codificatore funziona in modo inter-field od inter-frame; in tal caso infatti i coefficienti errati vengono antitrasformati generando campioni video errati su un intero blocco, che vengono memorizzati e usati come predizione per il semiquadro od il quadro successivo. Di conseguenza anche i nuovi campioni video ricostruiti risultano affetti da errore e la propagazione degli stessi continua finché il blocco non viene codificato in modo intra-field, cioè vengono trasmessi gli effettivi coefficienti trasformati.

4. Organizzazione a pacchetti

I sistemi di codifica basati sulla riduzione della ridondanza, ed in particolare quello qui descritto, sono caratterizzati da una elevata flessibilità di impiego. Infatti la loro adattatività alle caratteristiche del segnale da codificare e a quelle del canale di trasmissione ne permettono l'uso per applicazioni con obiettivi di qualità diversi e li rendono adatti a canali con capacità differenti.

Questo vantaggio intrinseco del sistema di codifica può essere ulteriormente ampliato da una struttura di trama video che garantisca: applicabilità a diversi formati televisivi e di trasmissione; facile recupero del sincronismo di trama, e possibilmente stretto legame fra la trama di trasmissione e quella video, allo scopo di facilitare le operazioni di sincronizzazione; elevata robustezza in caso di errori; limitata quantità di dati necessari per la gestione della trama stessa.

La struttura di trama video proposta è compatibile con diversi formati televisivi, in particolare i sistemi 4:2:2 e quelli HDTV attualmente allo studio. Tale compatibilità può essere ottenuta mantenendo inalterata l'organizzazione di base dei dati all'interno della trama, come risulterà evidente da una descrizione più dettagliata.

I dati video sono organizzati in pacchetti di lunghezza fissa, strutturalmente identici per tutte le applicazioni.

L'organizzazione di tipo comune consente anche un progetto di tipo modulare dei co-decodificatori; nel caso del segnale HDTV è necessario, con l'attuale tecnologia, elaborare in parallelo i dati video: la trama a pacchetti rende possibile l'adozione di una struttura modulare e parallela.

Il formato e le dimensioni scelti per i pacchetti sono adatti al tipo e caratteristiche dei dati, al codice di protezione dagli errori prescelto e conseguentemente permettono una elevata efficienza, riducendo al minimo la ridondanza richiesta per la gestione di trama.

Oltre ad adattarsi alle organizzazioni di trama per i vari bit-rate attualmente previsti dal CCITT, si può ritenere che l'attuale struttura a pacchetti permetta un facile adattamento alle reti di telecomunicazione ad accesso asincrono (ATM, Asynchronous Transfer Mode) attualmente allo studio a livello internazionale; tali reti permetteranno al segnale video di condividere il canale con altri servizi telematici e telefonici nell'ottica di creare una rete integrata.

4.1 CODICE A CORREZIONE DEGLI ERRORI E INTERLEAVING

Per proteggere i dati dagli errori distribuiti in modo uniforme è opportuno usare un codice a correzione. La scelta del codice deve rispondere congiuntamente all'esigenza di ottenere buone prestazioni, di introdurre una ridondanza non eccessiva, di essere realizzabile con circuiti integrati in grado di operare alle velocità richieste.

Tenendo presente queste esigenze una possibile scelta è rappresentata dai codici a blocco ciclici del tipo BCH (Bose, Chaudhuri, Hocquenghem). Questo tipo di codici, derivati dai codici di Hamming, presenta la massima efficienza in termini di rapporto fra capacità correttive introdotta. Per maggiori dettagli sulla struttura ed il funzionamento dei codici BCH in generale si veda la bibliografia 5.

Si è rivelato particolarmente adatto il codice BCH (255,239) che è in grado di correggere tutte le configurazioni di errore singole e doppie su parole di 255 bit con l'introduzione di soli 16 bit aggiuntivi e quindi di una ridondanza che risulta pari al 6,7%. È stato realizzato appositamente un circuito integrato in grado di effettuare la correzione nel caso di codici BCH (255,239) e BCH (511,493) e di quelli accorciati da essi derivati.

È già stato sottolineato come gli errori possono essere concentrati in burst; in tal caso un codice come quello proposto risulta inefficace. Si potrebbe ricorrere ad altri codici quali quelli di Reed-Solomon che sono in grado di correggere simboli, corrispondenti a sequenze di bit di lunghezza prefissata, all'interno delle parole codificate; essi risultano molto efficienti in presenza di burst di errori.

Una tecnica alternativa consiste nel ricorrere all'uso congiunto di un codice adatto alla correzione di errori uniformemente distribuiti, come il BCH, e di un fattore di interleaving (o interlacciamento): essa garantisce buone prestazioni sia in presenza di errori isolati che di burst.

Si è quindi adottato il codice BCH suddetto, più un fattore di interleaving pari a 16. I dati video vengono organizzati e protetti mediante 16 parole BCH come illustrato in figura 2.

Utilizzando questa tecnica, che non prevede l'introduzione di ridondanza aggiuntiva, il codice BCH è ancora in grado di correggere gli errori anche quando si verificano burst di lunghezza fino ad un massimo di 32 bit.

I pacchetti sono di lunghezza fissa e costituiti da 255 parole di 16 bit ciascuna per un totale di 4080 bit. Di queste, le ultime 16 contengono la ridondanza relativa al codice BCH.

Tutti i dati sono organizzati in pacchetti di tali dimensioni: si esamina ora nel dettaglio l'organizzazione dei dati video.

Fig. 2 - Organizzazione del codice BCH e fattore di interleaving. I dati vengono scritti e letti per colonne mentre la ridondanza del codice BCH viene calcolata sulle righe realizzando il fattore di interleaving richiesto.

4.2 STRUTTURA DEI PACCHETTI

Nel caso del sistema 4:2:2 si prevede che i dati video siano contenuti in due diversi tipi di pacchetti: i primi, detti di tipo F, contengono i dati a lunghezza fissa, mentre i secondi, detti di tipo V, contengono i dati a lunghezza variabile, cioè i coefficienti DCT codificati con parole a lunghezza variabile.

In Appendice viene descritta in maggiore dettaglio la struttura dell'intestazione (Header) mediante la quale vengono distinti i due tipi di pacchetto F e V e l'organizzazione dei pacchetti stessi, quale attualmente proposta per il sistema 4:2:2.

La distinzione dei due tipi di informazione in pacchetti diversi, identificati come due sorgenti distinte, semplifica la struttura del codificatore e del decodificatore, con vantaggi sensibili nella progettazione degli apparati. Infatti i due tipi di informazioni, F e V, sono prodotti a velocità diverse: le informazioni F rappresentano meno di 1 Mbit/s, mentre il restante flusso è occupato dalle informazioni V. È quindi possibile operare sui dati F con circuiti più complessi, per esempio programmabili, ma meno veloci, mentre è preferibile che i circuiti che operano sulle parole VLC siano semplici, organizzabili in modo parallelo, e molto più veloci.

Appare quindi evidente come le scelte relative a VLC, struttura di trama e strategie di protezione dagli errori, non siano indipendenti ed abbiano profonde implicazioni sulla fattibilità del co-decodificatore.

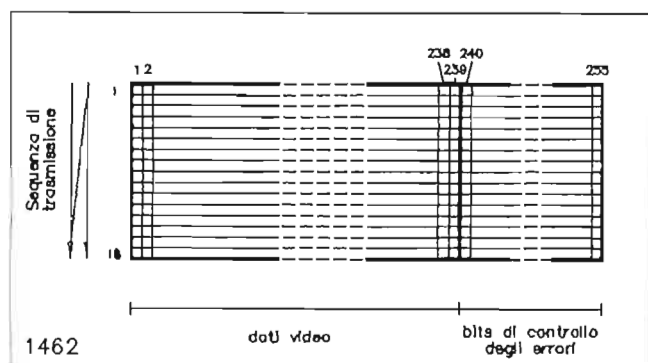
Ciò è particolarmente importante nel caso di codec capaci di operare sul segnale HDTV, dove la quantità di dati da elaborare nell'unità di tempo è quattro cinque volte maggiore che per il segnale 4:2:2.

Nel caso di HDTV, il codificatore è costituito da moduli; ciascun modulo elabora una porzione dell'immagine televisiva. Questa struttura parallela viene rispecchiata anche dalla struttura di trama. Ciascun modulo si comporta come una sorgente distinta di informazione e genera pacchetti identificati mediante un opportuno indirizzo di sorgente. Rimane inoltre valida la distinzione, per ciascuna sorgente, fra pacchetti F e pacchetti V.

4.3 EFFICIENZA DELLA STRUTTURA DI TRAMA

Un parametro di valutazione della bontà della trama è l'efficienza, intesa come rapporto fra il numero medio di bit utilizzato per la codifica dei coefficienti DCT ed il numero complessivo di bit trasmessi. Ovviamente questo valore è funzione del bit-rate di trasmissione. Si considerano aggiuntive le informazioni costituite dai pacchetti di tipo F, le intestazioni di pacchetto e gli End Of Block. Si considerano essenziali per l'ottenimento della massima qualità di codifica solo i dati relativi ai coefficienti DCT, ed in particolare quelli codificati con parole a lunghezza variabile.

Sono state effettuate valutazioni a 30 Mbit/s e 15 M bit/s, due velocità di trasmissione significative rispet-

tivamente nel caso di canali di trasmissione con qualità contributo e in quello di canali per distribuzione secondaria o diffusione.

In base alla suddetta definizione della efficienza σ si sono determinati i seguenti valori:

$\sigma = 0,84$ (30 M bit/s)
$\sigma = 0,75$ (15 M bit/s)

L'efficienza diminuisce al decrescere della velocità di trasmissione, poiché rimangono costanti i dati considerati aggiuntivi propri dei pacchetti di tipo F. È quindi importante progettare una struttura di trama che garantisca un elevato sfruttamento della capacità di canale per trasmettere i coefficienti DCT, soprattutto ai bassi bit rate, dove le implicazioni sulla qualità delle immagini ottenibile sono più evidenti.

I suddetti valori di efficienza sono stati confrontati con quelli ottenuti in base ad altre configurazioni di trama ed hanno dimostrato che la struttura a pacchetti proposta ha caratteristiche migliori anche sotto questo punto di vista.

5. Prestazioni della struttura proposta in presenza di errori sul canale

Per poter affermare che la struttura di trama descritta precedentemente garantisce una buona qualità del segnale riprodotto è anche necessario valutare le prestazioni in termini di errori corretti e rivelati. Infatti la presenza di errori residui ha effetti in genere molto visibili sull'immagine video decodificata.

5.1 Prestazioni del codice BCH

Il maccanismo fondamentale di protezione dagli errori è costituito dal codice BCH. Per valutare le sue prestazioni si può calcolare il BER all'uscita del decodificatore BCH in funzione del BER sul canale.

In Tabella 1 sono riportati i valori di probabilità di errore dopo la correzione, supponendo in entrata una distribuzione uniforme degli errori ed utilizzando alcune ipotesi semplificative per ricavare i risultati.

Questi risultati rappresentano il numero medio di errori presenti dopo aver effettuato la correzione; tuttavia anche supponendo che la distribuzione degli errori sul canale sia uniforme, non sarebbe tale la distribuzione degli errori residui in uscita dal correttore. Infatti il codice BCH utilizzato è caratterizzato da una distanza di Hamming pari a cinque, ovvero le parole di codice differiscono fra loro di almeno cinque bit. È quindi possibile rivelare fino a quattro errori, se il codice viene usato solo per la rivelazione, oppure correggere fino a due errori.

Se il codice viene usato per la correzione, come in questo caso, e se si verificano tre o più errori, nel tentativo di correggerli si decodifica una parola che, come minimo, ha distanza cinque da quella trasmessa.

Da ciò consegue che, dopo il correttore, ciascuna parola di codice e non è affetta da alcun errore, se in entrata gli errori erano in numero inferiore o uguale a due, oppure presenta un numero di errori uguale o maggiore di cinque, se in entrata gli errori erano tre o più.

Di questo fatto è necessario tenere conto valutando la distribuzione degli errori all'interno del pacchetto.

Tabella 1
PROBABILITÀ D'ERRORE DOPO CORREZIONE MEDIANTE CODICE BCH (255,239)

probabilità d'errore sul canale	probabilità d'errore dopo correzione FEC
10^{-2}	$9,2 \cdot 10^{-3}$
$5 \cdot 10^{-3}$	$2,7 \cdot 10^{-3}$
10^{-3}	$4,4 \cdot 10^{-5}$
$5 \cdot 10^{-4}$	$6,0 \cdot 10^{-6}$
10^{-4}	$5,2 \cdot 10^{-8}$
$5 \cdot 10^{-5}$	$6,6 \cdot 10^{-9}$
10^{-5}	$5,3 \cdot 10^{-11}$
$5 \cdot 10^{-6}$	$6,3 \cdot 10^{-12}$

5.2 Tecnica di mascheramento dell'errore

In precedenza si è sottolineato come gli errori residui producano un degradamento notevole dell'immagine decodificata poiché danno origine a blocchi fortemente scorrelati con quelli adiacenti e tale disturbo può propagarsi ai semiquadri video successivi.

Si è quindi ritenuto opportuno sviluppare una tecnica di mascheramento degli errori o *error concealment*, che riduca il degradamento visibile dovuto agli errori residui. Tale tecnica consiste nel sostituire il blocco od un gruppo di quadriblocchi affetti da errori con una predizione opportuna.

I valori di predizione sono comunque disponibili nel decodificatore nel caso di sistema ibrido: infatti durante il normale funzionamento del codec si utilizzano come predizioni i campioni video memorizzati appartenenti ai due semiquadri precedenti (modi inter-field e inter-frame).

I campioni ottenibili da dati affetti da errore possono quindi essere sostituiti da campioni video corrispondenti appartenenti al semiquadro precedente coposizionato o ottenuti da quelli appartenenti al semiquadro immediatamente precedente. È evidente che in presenza di movimento, la sostituzione comporta comunque un degradamento; infatti, a differenza del caso di decodifica secondo i normali modi inter in cui si utilizza l'errore di predizione trasmesso, in questo caso si sostituiscono semplicemente i campioni video ritenuti errati con i valori di predizione, generalmente diversi da quelli effettivi. Tuttavia, data l'elevata correlazione temporale esistente fra semiquadri adiacenti, statisticamente, si ottiene un vantaggio sensibile rispetto alla decodifica dei dati affetti da errore. Questa conclusione è stata ampiamente confermata dalle simulazioni effettuate.

Per quanto riguarda la scelta del predittore, l'uso del semiquadro precedente coposizionato presenta il vantaggio di non richiedere alcuna operazione di interpolazione ed è ovviamente ottimo nel caso di porzioni di immagine statiche; l'uso del semiquadro immediatamente precedente fornisce invece una predizione temporalmente più prossima, migliore nel caso di porzioni d'immagine dinamiche. Dalle simulazioni effettuate su alcune sequenze di immagini in movimento è risultato che questo secondo tipo di predizione è preferibile.

5.3 CODICI A RIVELAZIONE DEGLI ERRORI

Occorre rivelare gli errori in modo efficiente: la rivelazione deve avvenire dopo la correzione effettuata mediante il codice BCH e deve interessare sia la codifica VLC dei coefficienti DCT, sia la codifica dei modi di trasmissione, in quanto entrambi producono un degradamento visivo analogo sui quadriblocchi interessati dagli errori.

La presenza di errori residui è rivelabile ogni qual volta essa comporta la decodifica di una sequenza errata di EOB oppure quando un blocco risulta essere composto da più di 64 coefficienti.

La perdita di un pacchetto viene invece rivelata grazie alla presenza di un indice di continuità incrementato ogni volta che viene trasmesso un pacchetto ed azzerato all'inizio di ogni quadro.

Oltre a questi criteri, che consentono di determinare la presenza di particolari errori, si prevede di aggiungere una limitata ridondanza con lo scopo specifico di permettere la rivelazione della quasi totalità degli errori residui.

Si propone l'adozione di codici ciclici usati solo per rivelare gli errori e detti usualmente CRC. La rivelazione dell'errore avviene tutte le volte che la sindrome calcolata in ricezione è diversa da zero (bibl. 6 e 7).

La scelta effettuata per il codice deriva da un compromesso fra prestazioni in termini di probabilità di rivelazione e ridondanza introdotta. In particolare si è voluto introdurre tali codici senza modificare in modo essenziale la struttura di trama e a tale scopo si sono utilizzati i pacchetti di tipo F organizzandoli nel modo descritto in Appendice.

Sono utilizzati due codici CRC: uno protegge le informazioni relative ai modi di trasmissione per un gruppo di otto quadriblocchi consecutivi, l'altro protegge le parole a lunghezza variabile corrispondenti allo stesso gruppo di quadriblocchi.

Per la protezione dei modi di trasmissione è stato scelto un codice (63,57) con il seguente polinomio generatore:

$$g_1(x) = x^6 + x^5 + 1$$

Il codice (63,57) opera come un codice accorciato in quanto la ridondanza è valutata su 48 bit di informazione.

Per la protezione dei dati a lunghezza variabile è stato scelto un codice che introduce 10 bit di ridondanza con il seguente polinomio generatore:

$$g_2(x) = x^{10} + x^7 + 1$$

Tale codice è applicato su un numero di bit variabile. Infatti mediamente il numero di bit codificati appartenenti a 8 quadriblocchi corrisponde a circa 3000 bit a 30 M bit/s e a circa 1500 bit a 15 M bit/s. Si è quindi in presenza di un utilizzo atipico del codice CRC poiché esso opera su un numero di bit superiore a quello convenzionale per codici che introducono 10 bit di ridondanza, cioè 1013 bit. Non è quindi possibile effettuare valutazioni teoriche sulle prestazioni del codice stesso; si può comunque presupporre un peggioramento delle prestazioni rispetto a quelle ottenibili per un uso convenzionale dovuto al fatto che alla sindrome composta da soli 0 corrispondono un numero maggiore di configurazioni di errore.

Per conoscere le prestazioni effettive di questi due codici è stato necessario ricorrere a simulazioni; affinché queste potessero applicarsi anche ai casi più sfavorevoli, si è simulata la distribuzione degli errori all'interno di un pacchetto tenendo conto della presenza del codice BCH e della propagazione degli errori dovuta ad una modulazione numerica con codifica differenziale in fase, ad esempio di tipo 4-PSK, ed alla presenza di scrambler autosincronizzanti.

I risultati delle simulazioni, effettuate per entrambi i codici, hanno mostrato che nel caso del codice (63,57) la probabilità di non rivelazione è molto inferiore al limite superiore teorico, che è pari a $1{,}5 \cdot 10^{-2}$, e vale circa $2 \cdot 10^{-5}$, mentre nel caso del codice che introduce 10 bit di ridondanza non ci si discosta molto dal limite stesso pari a 10^{-3}. Complessivamente le prestazioni dei due codici risultano adeguate.

5.4 STRATEGIE DI RINFRESCO

Benché l'uso del mascheramento degli errori comporti in generale un miglioramento percepibile della qualità delle immagini riprodotte anche in presenza di elevati tassi di errore, non bisogna sottovalutare i degradamenti comunque presenti nelle aree soggette a «error concealment» caratterizzate da movimento e da elevato dettaglio.

Infatti nel caso in cui la porzione di immagine, in genere un gruppo di 8 quadriblocchi, soggetta ad errori rivelati viene «mascherata», i pel ad essa appartenenti vengono memorizzati nella memoria di quadro e utilizzati come predizioni nei semiquadri successivi, se il funzionamento in modo inter del codificatore lo impone. Si ha quindi una propagazione del degradamento nei quadri successivi con una conseguente maggiore visibilità dello stesso.

Fortunatamente sono proprio le aree più critiche per il mascheramento, cioè quelle dinamiche, quelle per cui viene normalmente impiegato il modo intrafield.

È comunque conveniente forzare periodicamente il codificatore ad operare in modo intra-field così da bloccare la propagazione degli errori e da operare un rinfresco o *refresh* dell'informazione presente in ricezione.

La tecnica di rinfresco dell'immagine comporta un incremento del numero medio dei bit di codifica per pel e quindi è necessario anche in questo caso trovare un compromesso tra incremento della ridondanza e riduzione del periodo di rinfresco.

Fra le possibili strategie di rinfresco ne sono state studiate due. La prima consiste nel forzare periodicamente un intero quadro in modo intra-field. La seconda invece forza in modo intra-field i singoli quadriblocchi oppure i gruppi di 8 quadriblocchi se ciascuno di questi ha operato in modo inter-field o interframe consecutivamente per un numero di semiquadri superiore ad un massimo prefissato.

La seconda tecnica risulta vantaggiosa in termini di minore incremento della ridondanza in quanto tiene conto del funzionamento in modo intra-field dovuto a sole ragioni di codifica ottima e sfrutta il rinfresco così ottenuto in modo «naturale». D'altro canto questa seconda tecnica richiede la presenza di un elevato numero di contatori nel codificatore (3240 per i singoli quadriblocchi e 405 nel caso in cui si raggruppino gruppi di 8 quadriblocchi) per mantenere memoria dei modi di trasmissione (inter-field o inter-frame) usati nel tempo.

5.5 PRESTAZIONI COMPLESSIVE DEL SISTEMA IN PRESENZA DI ERRORI

Tenendo complessivamente conto della correzione degli errori effettuata dal codice BCH, della distribuzione

Tabella 2.
RIVELAZIONE DEGLI ERRORI RESIDUI

tasso di errore sul canale	gruppi di 8 quadriblocchi (g) affetti da errori			
	30 Mbit/s		15 Mbit/s	
	rivelati	non rivelati	rivelati	non rivelati
10^{-3}	390 g/s	18,4 g/min.	313 g/s	10,1 g/min.
$5 \cdot 10^{-4}$	54 g/s	2,5 g/min.	43,8 g/s	1,4 g/min.
10^{-4}	27,8 g/min.	1,3 g/ora	22,5 g/min.	0,7 g/ora
$5 \cdot 10^{-5}$	3,6 g/min	4 g/giorno	2,8 g/min.	2,2 g/giorno
10-5	1,7 g/ora	1 g/mese	1,4 g/ora	0,5 g/giorno
$5 \cdot 10^{-6}$	3,9 g/giorno		2,8 g/giorno	

Nota: valutazione teorica del numero di gruppi di 8 quadriblocchi affetti da errore rivelati e non rivelati a due differenti velocità di trasmissione (30 e 15 Mbit/s). I dati sono ottenuti nel caso in cui si utilizzi codice BCH (255,239) per la correzione degli errori e due CRC per la rivelazione degli errori residui. La tecnica del mascheramento degli errori può inoltre essere applicata per i gruppi di 8 quadriblocchi (un area di immagine pari a 128 · 8 pel) per i quali viene rivelata la presenza di errori residui.

degli errori residui dopo la correzione e dell'effetto della tecnica di mascheramento mediante i codici CRC descritti si può giungere ad una valutazione globale delle prestazioni del sistema in termini di numero di gruppi di 8 quadriblocchi errati e mascherabili con la tecnica di error concealment.

I risultati valutati teoricamente a 30 Mbit/s e 15 Mbit/s sono riportati nella Tabella 2.

Si può notare una notevole riduzione del numero degli errori residui non rivelati.

L'effettivo vantaggio qualitativo per l'immagine riprodotta è di difficile valutazione, poiché è determinabile solo su base statistica, dipendendo dalla localizzazione degli errori e dal contenuto dell'immagine stessa. Si sono effettuate simulazioni con calcolatore su due sequenze ipotizzando diverse probabilità di errore e con una distribuzione degli errori ritenuta significativa.

Estrapolando questi risultati parziali si può ritenere che gli effetti visibili degli errori non siano fastidiosi per BER sul canale minori o uguali a 10^{-4} e che si ottenga comunque un sensibile miglioramento anche a BER superiori.

6. Conclusioni

Normalmente i sistemi di codifica e di trasmissione di tipo analogico presentano un degradamento della qualità del segnale riprodotto che aumenta gradualmente al crescere del rumore e delle distorsioni introdotte dal canale. I sistemi numerici sono invece caratterizzati da una qualità di base migliore e da un suo peggioramento brusco quando si raggiunge la soglia operativa del servizio. Questo comportamento dei sistemi numerici può essere reso meno brusco grazie ad un accurato progetto dell'organizzazione dei dati trasmessi. In particolare la tecnica di mascheramento degli errori proposta, consente un aumento graduale della visibilità degli errori al crescere della probabilità di errore sul canale.

L'esperienza acquisita nello sviluppo di un sistema di codifica basato su DCT puramente intra-field e che ha portato alla realizzazione di un co-decodificatore progettato, prodotto e commercializzato dalla Telettra S.p.A., ha permesso di focalizzare l'estrema importanza della struttura dei dati in trasmissione sulle prestazioni e sulla qualità finale ottenibili.

Una scelta opportuna di codice a lunghezza variabile, di struttura di trama video e di schemi di protezione dagli errori consente di sfruttare al massimo la flessibilità e l'efficienza dei sistemi di riduzione della ridondanza basati su DCT, di conseguenza ne consente l'impiego per diverse applicazioni, con canali aventi diverse caratteristiche, e per un'ampia gamma di velocità di trasmissione.

La struttura di trama proposta, basata su pacchetti di lunghezza fissa, presenta ottime caratteristiche dal punto di vista della facilità di sincronizzazione e della robustezza in caso di errori.

Ringraziamenti

Fondamentale per la definizione della proposta di organizzazione dei dati descritta in questo articolo è stato il contributo di idee dell'ing. Silvio Cucchi della Telettra S.p.A.

APPENDICE

TRAMA A PACCHETTI

Ciascun pacchetto è costituito da 255 parole di 16 bit. La prima parola è l'intestazione o *header* contenente le seguenti informazioni:

— Identificatore del servizio SI (3 bit): identifica il tipo di segnale contenuto cioè, ad esempio, se si tratta del segnale video, del segnale audio, di segnali ausiliari oppure altri servizi, ad esempio di Televideo.

— Indirizzo della sorgente SA (4 bit): nel caso in cui ad un singolo servizio appartengano pacchetti di tipo diverso, essi vengono contraddistinti come generati da sorgenti diverse.

— Indice di continuità CI (3 bit): è un contatore che viene incrementato ad ogni invio di pacchetto e permette, in ricezione, di controllare la perdita di pacchetti.

L'incremento avviene in modo ciclico ed il contatore viene azzerato all'inizio di ogni quadro.

— Indice di pacchetto completo o incompleto CF (1 bit): permette di indicare se il pacchetto è costituito da soli dati utili oppure se il pacchetto è incompleto e parte dei dati sono di riempimento. Ad esempio, nel caso dell'ultimo pacchetto appartenente ad un quadro video, esso può essere incompleto.

I restanti 5 bit (EP) della parola vengono usati per un codice di Hamming (15,11) che consente la correzione degli errori singoli sui precedenti 11 bit più un bit di parità totale che permette la rivelazione degli errori di ordine pari. Queste protezioni aggiuntive sono necessarie data l'importanza dei dati contenuti nell'header.

Nel caso dello standard 4:2:2 vi sono due tipi di pacchetti contenenti i dati video: i pacchetti di tipo F contengono i dati a lunghezza fissa, i pacchetti di tipo V contengono la codifica VLC dei coefficienti DCT.

Pacchetti di tipo F

I pacchetti di tipo F contengono le informazioni a lunghezza fissa. Nel caso dello standard 4:2:2 a 625 linee vengono trasmessi 8 pacchetti di tipo F per ogni quadro.

Le 239 parole disponibili per ogni pacchetto sono così suddivise:

0	Header
1	Dati ausiliari
2 ÷ 25	Componenti dei vettori movimento
26 ÷ 34	Fattori di trasmissione
35 ÷ 238	204 parole suddivise in 51 gruppi, ciascuno relativo ad un gruppo di 8 quadriblocchi con l'indicazione dei modi di trasmissione relativi.

La struttura dei pacchetti di tipo F è schematizzata in figura 3.

La parola di dati ausiliari contiene 3 bit per l'identificazione del formato video (ST), 1 bit che indica se si tratta di dati relativi ad un semiquadro pari o dispari (FS),

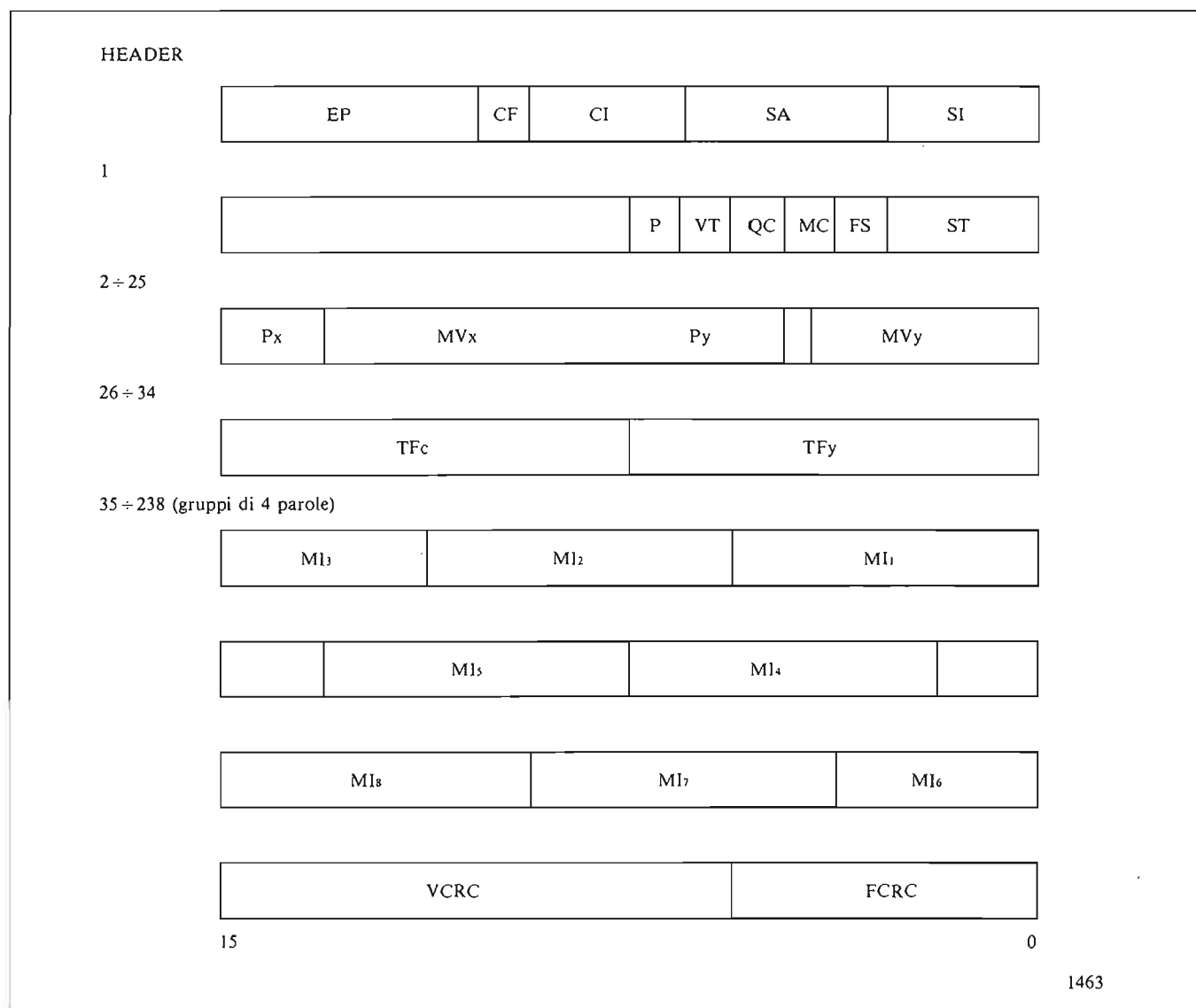



Fig. 3 - Struttura dei pacchetti di tipo F.

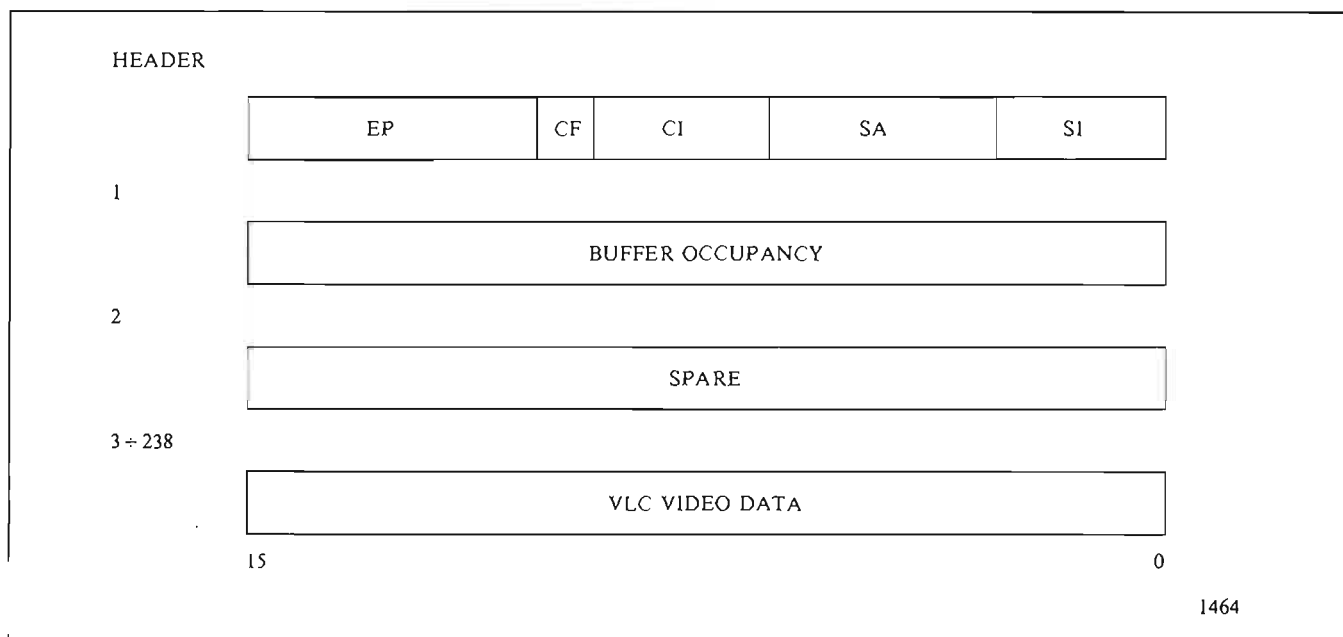



Fig. 4 - Struttura dei pacchetti di tipo V.

1 bit relativo all'uso della compensazione del movimento (MC), 1 bit che indica se viene usata la quantizzazione non lineare nella codifica intrafield della luminanza (QC), 1 bit che indica se viene applicata o no la matrice di visibilità (VT), ed infine 1 bit di parità (P).

Nel caso in cui siano usati i 32 vettori movimento è necessario, per ciascuno di essi, indicarne le due componenti (MV_x e MV_y); queste sono protette rispettivamente da 2 e 3 bit di parità (P_x e P_y). Le componenti sono contenute nelle 24 parole disponibili per ogni pacchetto e vengono ripetute 3 volte all'interno dei 4 pacchetti disponibili per ogni semiquadro secondo il seguente ordine:

Pacchetto 0: vettori $0 \div 7$ $8 \div 15$ $16 \div 23$
Pacchetto 1: vettori $8 \div 15$ $16 \div 23$ $24 \div 31$
Pacchetto 2: vettori $16 \div 23$ $24 \div 31$ $0 \div 7$
Pacchetto 3: vettori $24 \div 31$ $0 \div 7$ $8 \div 15$

la ripetizione dei vettori movimento permette di effettuare una decisione a maggioranza, nel caso in cui siano presenti degli errori residui dopo la correzione effettuata dal codice BCH.

I fattori di trasmissione inviati sono relativi ad ogni striscia ed occupano ciascuno una parola, indicando, sui due byte disponibili, il fattore di trasmissione della luminanza (TF_y) e della crominanza (TF_c) rispettivamente.

L'ultimo gruppo di parole prevede la codifica dei modi di trasmissione. Questa è effettuata mediante 6 bit (MI) codificando, secondo un'opportuna tabella, il modo di trasmissione relativo ad ogni singolo quadriblocco oppure, se viene usata la compensazione movimento, fornendo l'indirizzo del vettore movimento, fra i 32 disponibili, da utilizzare per quel quadriblocco. Le parole di 6 bit sono raggruppate in gruppi di 8, occupando così 3 parole di 16 bit a cui ne va aggiunta una quarta che contiene 6 bit di ridondanza per il codice (63,57) (FCRC), usato per rivelare la presenza di errori residui sui precedenti 48 bit, ed i rimanenti 10 bit che rappresentano la ridondanza relativa alla codifica dei coefficienti DCT dei corrispondenti 8 quadriblocchi (VCRC).

Pacchetti di tipo V

I pacchetti di tipo V contengono essenzialmente i coefficienti DCT codificati. Come mostra la figura 4, la prima parola dopo l'header contiene l'occupazione del buffer, espressa in numero di pacchetti, mentre la parola successiva non contiene, attualmente, alcun tipo di dati. Le rimanenti 236 parole contengono i coefficienti DCT (dati video VLC).

Come è evidente, il numero di pacchetti di tipo V non è fisso ma dipende, per la presenza di un codice VLC, dal bit-rate di trasmissione e dal contenuto informativo della porzione di immagine trasmessa.

(3845)

BIBLIOGRAFIA

1. - BARBERO M., STROPPIANA M.: *Codifica del segnale televisivo numerico: sistemi di riduzione della ridondanza mediante l'uso della trasformata coseno discreta.* «Elettronica e Telecomunicazioni», n. 1, 1989, pagg. 17-29.

2. - IWP CMTT/2: *Revised version of Draft Report AD/CMTT: Digital Transmission of Component-Coded Television Signal at 30-34 Mbit/s and 45 Mbit/s.* Giugno 1989.

3. - BARBERO M., BELLORA R., STROPPIANA M.: *Codifica del segnale video numerico: Codici a lunghezza variabile applicati ai coefficienti DCT.* «Elettronica e Telecomunicazioni», in questo numero.

4. - IWP 11/7: *Draft Report AG/11 User Requirements for Digital Television Transmission.* Giugno 1989.

5. - MAC-WILLIAMS F.J., SLOANE N.J.A.: *The theory of error correcting codes.* North Holland 1978.

6. - PETERSON W.W., BROWN D.T.: *Cyclic codes for error detection.* Proceedings of the IRE, gennaio 1961.

7. - HELLMAN M.E.: *Error detection made simple.* Department of Electrical Engineering - Stanford University.

CODIFICA DEL SEGNALE TELEVISIVO NUMERICO: CODICI A LUNGHEZZA VARIABILE APPLICATI AI COEFFICIENTI DCT

M. BARBERO, R. BELLORA, M. STROPPIANA (*)

SOMMARIO — Gli algoritmi per la riduzione della ridondanza del segnale televisivo sono oggetto di studio presso il Centro Ricerche della RAI. Una tecnica di riduzione che si è dimostrata particolarmente efficiente fa uso della Trasformata Coseno Discreta (DCT) e della codifica dei coefficienti trasformati mediante codici a lunghezza variabile (VLC). Tale tecnica è utilizzata per la realizzazione di codificatori per la trasmissione di segnali televisivi a definizione normale e ad alta definizione (HDTV). In questo articolo si descrivono in dettaglio i codici a lunghezza variabile. Dopo una panoramica su alcuni codici fra i più diffusi, sono analizzati quelli che sono particolarmente adatti per la codifica dei coefficienti DCT ottenuti da segnale video e che per tale ragione sono stati utilizzati in pratica e sono attualmente alla base delle proposte di normalizzazione in discussione presso i gruppi internazionali. È disponibile la versione inglese del presente articolo.

SUMMARY — ***Digital TV signal coding: Variable Length Codes applied to DCT coefficients:*** *The algorithms for the bit-rate reduction of the television signal are studied at the RAI Research Centre. A technique which is particularly efficient makes use of the Discrete Cosine Transform (DCT) and of the coding of the transform coefficients by a Variable Length Code (VLC). Such technique is used to implement codecs for the transmission of standard definition TV signal and High Definition TV signals (HDTV). In this article the VLCs are described in details. After an overview on some of the most employed VLCs, those especially suitable for coding the DCT coefficients derived from the video signal are analyzed; they have been practically implemented and are the basis of proposals for standardisation under discussion in the international groups. The English version of this article is available.*

1. Introduzione

In articoli precedenti (bibl. 1, 2) si è affrontato l'argomento della codifica numerica del segnale televisivo e si sono analizzati i metodi della riduzione della ridondanza al fine di limitare il flusso di dati necessario a trasmettere o memorizzare tale segnale, garantendo comunque una buona qualità dell'immagine.

Si è determinato che l'algoritmo di Trasformata Discreta Coseno (DCT, Discrete Cosine Transform) presenta un'elevata efficienza. Sono stati realizzati sistemi utilizzanti tale algoritmo; se di tipo intra-field, essi operano sugli elementi di immagine appartenenti ad un semiquadro televisivo ed eliminano solamente la ridondanza spaziale; di tale tipo è un codificatore progettato e realizzato in Italia (bibl. 3). Sistemi più sofisticati sono di tipo Tridimensionale o Ibrido, essi eliminano anche la ridondanza temporale (bibl. 2).

In tali sistemi la trasformata DCT viene effettuata su *blocchi* di $8 \cdot 8$ elementi di immagine (pel) fornendo coefficienti scorrelati tra loro. I coefficienti DCT presentano una distribuzione non uniforme delle ampiezze e quindi è conveniente, al fine di ridurre il flusso binario dei dati, codificarli mediante Codici a Lunghezza Variabile (VLC, Variabile Length Code) che associano ai simboli dell'alfabeto della sorgente parole binarie di diversa lunghezza. Esistono vari tipi di VLC, con caratteristiche differenti fra loro e pertanto adatti ad applicazioni diverse.

Nella prima parte dell'articolo si definiscono i concetti di entropia e di codifica dell'alfabeto di una sorgente.

Successivamente si analizzano alcuni tipi di VLC con particolare attenzione a quelli oggetto di studio nel Gruppo di lavoro internazionale CMTT/2; tale gruppo ha il compito di definire un sistema di codifica per la trasmissione del segnale televisivo 4:2:2 conforme alla Raccomandazione 601 del CCIR. Infine si riportano i risultati relativi all'ottimizzazione del codice B_2 che è uno dei VLC allo studio presso il CMTT/2 (bibl. 4) ed è utilizzato per la realizzazione di sistemi sviluppati in Italia (bibl. 3, 5).

Il presente lavoro è stato svolto nell'ambito del progetto europeo EUREKA 256.

2. Generalità sui VLC

La codifica a lunghezza variabile dell'alfabeto di una sorgente consiste nell'assegnare, ad ogni simbolo generato, una parola di codice la cui lunghezza, in numero di bit, è inversamente proporzionale alla sua probabilità. In tale modo si riesce a ridurre il numero medio di bit per

(*) Ing. Marzio Barbero, dott. Mario Stroppiana del Centro Ricerche RAI - Torino; ing. Roberto Bellora del Politecnico di Torino.
Dattiloscritto pervenuto alla redazione l'8 agosto 1989.

simbolo; i simboli più frequenti sono codificati con pochi bit, mentre quelli meno frequenti sono codificati con un numero maggiore di bit.

Il codice deve inoltre essere tale da consentire una esatta decodifica da parte del ricevitore. A tal fine esso deve soddisfare la *regola del prefisso*; essa afferma che nessuna parola di codice deve essere uguale al prefisso di un'altra.

Per comprendere come tale tipo di codifica minimizza il numero medio di bit per simbolo, si fa riferimento ai concetti di *entropia dell'alfabeto di sorgente* e di *informazione* associata ad ogni suo simbolo e si valuta l'*efficienza del codice*, definita come rapporto tra l'entropia ed il numero di bit/simbolo.

2.1 Entropia di sorgente

Si consideri una sorgente discreta che genera un alfabeto A costituito da M simboli a_j aventi, ognuno, probabilità $p(a_j)$ di essere generato e quindi sia $p(a_j) \geq 0$ e $\sum_j p(a_j) = 1$. Ad un simbolo con bassa probabilità è associato un contenuto di informazione superiore rispetto a quello dei simboli più frequenti. Si definisce quindi *Informazione* $I(a_j)$, espressa in bit, di un generico simbolo a_j di una sorgente discreta, come:

$$I(a_j) = \log_2 (1/p(a_j)) \quad [1]$$

È definita l'*entropia dell'alfabeto di sorgente* H(A) come:

$$H(A) = \sum_{j=1}^{M} p(a_j) \cdot I(a_j) \quad [2]$$

Se l'informazione è data dalla [1], ossia è espressa in bit, allora l'entropia rappresenta il minimo numero medio di bit necessari a codificare i simboli dell'alfabeto A e tale valore soddisfa la relazione:

$$0 \leq H(A) \leq \log_2(M) \quad [3]$$

dove $\log_2(M)$ indica l'entropia di un alfabeto in cui i simboli a_j hanno tutti la stessa probabilità $p(a_j) = 1/M$.

L'entropia H(A) dell'alfabeto risulta uguale all'entropia della sorgente solo nel caso in cui i simboli siano indipendenti tra loro; in caso contrario l'entropia della sorgente che è data dall'entropia dei suoi simboli, noti i simboli precedenti, risulta inferiore a quella dell'alfabeto (bibl. 6).

Ci si può avvicinare all'entropia di sorgente codificando, invece dei singoli simboli, un loro raggruppamento; si ottiene in tal modo quella che viene normalmente indicata come quantizzazione vettoriale (bibl. 7).

2.2 Codifica di sorgente

Si supponga di voler trasmettere i simboli dell'alfabeto di una sorgente su un canale binario; ogni simbolo deve essere codificato con un numero intero di bit. La *lunghezza media* delle parole di codice risulta quindi:

$$\tilde{n} = \sum_{j=1}^{M} p(a_j) \cdot n_j \quad [4]$$

dove M indica il numero di simboli dell'alfabeto A ed n_j il numero di bit della parola di codice associata al simbolo a_j. Scopo della codifica è quello di minimizzare tale grandezza. L'informazione associata ad ogni simbolo può essere, al massimo, uguale al numero di bit n_j con cui esso è codificato. I valori n_j sono quindi maggiori od uguali all'informazione $I(a_j)$ associata al simbolo a_j e quindi la lunghezza media delle parole di codice è maggiore od al limite uguale all'entropia dell'alfabeto. Inoltre si può dimostrare che esiste almeno un codice, soddisfacente la regola del prefisso, tale che:

$$H(A) \leq \tilde{n} \leq H(A) + 1 \quad [5]$$

La lunghezza media $\tilde{n}$ risulta uguale all'entropia dell'alfabeto solo nel caso in cui la probabilità di ogni simbolo sia data da:

$$p(a_j) = 2^{-n_j} \quad [6]$$

Per ogni codice si valuta un valore di efficienza α dato dal rapporto tra l'entropia dell'alfabeto da codificare e la lunghezza media delle parole del codice, ossia:

$$\alpha = \frac{H(A)}{\tilde{n}} \quad [7]$$

Si noti che la codifica a lunghezza variabile genera un flusso di bit non costante.

3. **Alcuni tipi di VLC**

I codici descritti soddisfano tutti la regola del prefisso e quindi sono univocamente decodificabili.

Il codice di Huffman è costruito, come si può desumere dall'esempio 3.2.1, in base alla statistica dell'alfabeto da codificare e quindi risulta il codice ottimo. Questa massima libertà nella costruzione del codice presenta però alcuni inconvenienti che saranno meglio specificati nel seguito.

La determinazione degli altri codici avviene invece in due fasi; nella prima fase si individua una classe di codici e nella seconda fase, in base alla statistica dell'alfabeto, si stabilisce il codice appartenente alla classe scelta e si assegnano le parole di codice ai simboli.

La libertà concessa nella determinazione di tali codici risulta quindi inferiore rispetto a quella permessa nella costruzione del codice di Huffman. Questi, chiamati codici subottimi, hanno un'efficienza inferiore, ma presentano altri vantaggi, ad esempio semplicità di codifica e comportamento migliore in presenza di errori.

Si fa notare che la libertà nella costruzione del codice può essere ulteriormente limitata individuando, nella prima fase, un solo codice, invece di una classe di codici, ed utilizzando, nella seconda fase, la statistica dell'alfabeto per la sola assegnazione delle parole. Con questa ulteriore restrizione, in linea generale si riduce ulteriormente l'efficienza del codice, ma si accentuano gli altri vantaggi, come la semplicità di co-decodifica.

Prima di descrivere alcuni VLC si illustra la codifica di sequenze di simboli uguali (RLC, Run Length Coding).

3.1 Codifica di Sequenze di simboli uguali (RLC)

La codifica di sequenze di simboli uguali consente una precodifica dell'alfabeto di sorgente. Il nuovo alfabeto è

ottenuto raggruppando sequenze di simboli consecutivi uguali ed assegnando a loro un simbolo del nuovo alfabeto. Le parole di questo codice intermedio possono essere codificate con codici a lunghezza fissa o con codici a lunghezza variabile.

Questo tipo di precodifica risulta efficiente quando la probabilità di avere n simboli uguali consecutivi è superiore al prodotto di n volte la probabilità di occorrenza di quel simbolo.

Un esempio in cui si verifica tale situazione è quello della trasmissione in facsimile (immagini a due livelli) dove l'alfabeto di sorgente è costituito dai soli simboli binari «0» ed «1» e sono molto frequenti le sequenze di bit «0» e di bit «1» in quanto esiste una forte correlazione tra bit adiacenti.

Sono illustrati due esempi di uso degli RLC applicati alla codifica del segnale televisivo mediante DCT.

3.1.1 *Esempio: uso del RLC per la codifica di coefficienti DCT*

Nel caso del segnale video è stato proposto il seguente metodo per codificare i coefficienti DCT (bibl. 8).

Si considerano i coefficienti appartenenti ad un insieme di blocchi DCT di 8 · 8 pel e si codificano i coefficienti dello stesso ordine di tutto l'insieme sfruttando la procedura descritta in figura 1.

Questo sistema di codifica presenta un'elevata efficienza, analoga a quella ottenibile mediante l'uso di VLC che sono descritti nel seguito.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
3	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.
4	.	.	1	.	.	.	1	.	.	0	.	.	.	.	.	.	1	.	.	.
5	.	.	1	.	1	.	0	.	.	0	.	.	.	.	.	.	1	1	.	.
6	.	.	1	.	0	.	1	.	.	1	.	.	.	.	.	1	0	0	1	.
7	.	1	0	.	0	.	1	.	.	0	.	.	.	.	.	0	1	1	1	.
8	.	0	1	1	0	.	1	.	.	0	.	.	.	.	.	0	1	0	1	.
9	1	0	1	0	1	1	1	.	.	0	.	.	.	.	.	0	0	0	0	.
10	0	0	1	0	0	1	1	1	.	0	.	1	.	1	.	0	1	0	1	1
11	0	1	0	1	1	1	1	0	1	1	.	0	1	1	.	1	1	0	0	1
12	0	1	1	1	0	1	1	0	0	1	1	1	0	1	1	1	0	0	0	1

1473

Fig. 1 — Esempio di uso di RLC per codificare i coefficienti DCT. Tutti i coefficienti omologhi, dello stesso ordine, appartenenti ad un gruppo di blocchi, nell'esempio 20 blocchi, sono organizzati in una tabella come in figura, dove in ordinate sono disposti i bit appartenenti a ciascun coefficiente, in posizione 1 è il bit più significativo e in posizione 12 quello meno significativo. A partire dalla posizione (1,1) si procede a codificare con un RLC tutti gli zeri più significativi procedendo da sinistra a destra, dall'alto al basso; dopo 29 zeri si giunge a (10,3) che è il primo bit diverso da zero; si trasmette il coefficiente di colonna 10, ovvero 1001000011, e si azzerano le posizioni da esso occupate; si procede quindi a codificare mediante RLC la sequenza compresa fra (11,3) e (2,4) per poi trasmettere il coefficiente 3, ovvero 111011101, e azzerare le posizioni corrispondenti. Allo stesso modo si codificano tutti i 20 coefficienti considerati.

3.1.2 *Esempio: codifica delle sequenze di coefficienti DCT uguali a zero*

In questo caso l'uso del RLC è di ausilio a quello del VLC per aumentarne l'efficienza. Nel caso dei coefficienti DCT del segnale televisivo sono molto frequenti le sequenze di simboli uguali a zero. Questo è dovuto al fatto che la DCT concentra l'energia sui coefficienti di bassa frequenza (bibl. 9). I coefficienti DCT di ordine elevato hanno così un'energia molto bassa e quindi è molto probabile, per loro, il valore zero. Ordinando quindi i coefficienti DCT per mezzo di un opportuno percorso di scansione, ad esempio a zigzag, si rendono consecutivi i coefficienti di ordine elevato.

È quindi possibile utilizzare un RLC per codificare i coefficienti nulli consecutivi, con vantaggi sensibili rispetto alla loro codifica separata.

Un ulteriore vantaggio è ottenibile mediante l'uso di un simbolo di fine blocco (EOB, End Of Block) che consente di evitare la trasmissione della sequenza di zeri che si trova alla fine del blocco. In tale modo si concentrano tutte le ultime sequenze di zeri in un unico simbolo la cui frequenza risulta elevata ed è perciò codificato con una parola di pochi bit, aumentando in tal modo l'efficienza di codifica.

3.2 CODICE DI HUFFMAN

Il codice di Huffman minimizza la lunghezza media delle parole di codice in base alla statistica dell'alfabeto della sorgente e quindi risulta il codice ottimale.

Ciò appare evidente dall'esame del modo in cui tale codice è costruito: un esempio della procedura per la codifica di un alfabeto discreto di simboli (bibl. 10) è riportata qui di seguito; una descrizione completa della procedura si trova in bibl. 11.

È evidente che al variare della statistica dell'alfabeto si ottengono differenti codici di Huffman e che tale codice si allontana dalle condizioni di ottimalità quando codifica simboli di una sorgente avente una statistica diversa da quella utilizzata per la determinazione del codice.

Riferendoci sempre al modo di generazione del codice, si possono individuare alcuni svantaggi legati al suo utilizzo.

La suddivisione del messaggio nelle singole parole è legata al riconoscimento delle stesse, quindi la sincronizzazione di parola non può essere disgiunta dalla decodifica.

Gli effetti degli errori introdotti dal canale di trasmissione non si limitano alla sola parola affetta da errori, ma si propagano anche alle parole successive. Infatti un errore può trasformare una parola di codice in un'altra, più lunga oppure più corta, e questo può impedire la corretta decodifica delle parole successive.

3.2.1 *Esempio: un algoritmo di costruzione del codice di Huffman*

Supponiamo di disporre di una sorgente discreta costituita da un alfabeto di simboli e di volere generare il codice di Huffman corrispondente.

Fig. 2 — Esempio di costruzione del codice di Huffman. In questo semplice caso si ha un alfabeto di 6 simboli; il procedimento per la costruzione del codice richiede innanzitutto l'ordinamento dei simboli in entrata in base alla loro probabilità di occorrenza (per i simboli equiprobabili l'ordine non ha importanza). A questo punto le due probabilità più piccole, nell'esempio 0,06 e 0,04, sono sommate fra loro e viene costituito un nuovo insieme ordinato delle probabilità (primo passo). Questa operazione viene poi ripetuta fino a quando si hanno solo più due probabilità complementari (quarto passo).

In figura 2 è indicata la procedura che può essere seguita per costruire il codice relativo ad una sorgente costituita da 6 simboli. L'assegnazione delle parole di codice ai simboli è poi effettuata mediante la procedura descritta in figura 3.

L'entropia della sorgente considerata nell'esempio può essere calcolata secondo la [2] ottenendo H(A) = 2,14 bit, mentre la lunghezza media della parola di codice secondo la [4] è $\bar{n}$ = 2,2 bit; in questo esempio il rendimento ottenuto è α = 97,3%.

3.3 COMMA CODE

Il Comma code è un codice subottimo che diventa ottimo, coincidente con il codice di Huffman, quando i simboli dell'alfabeto da codificare hanno una probabilità di occorrenza data dalla [6] con $n_j = j$. Le parole di codice si differenziano, una dall'altra, in base al numero di ripetizione di un determinato bit, mentre l'altro bit funge da separatore tra le parole, ovvero da virgola o «comma» in inglese; un esempio è riportato in figura 4.

a_j	$p(a_j)$	comma	Huffman
1	0,4	0	0
2	0,175	10	111
3	0,175	110	110
4	0,09	1110	100
5	0,09	11110	1011
6	0,028	111110	10100
7	0,028	1111110	101011
8	0,005	11111110	1010100
9	0,005	111111110	10101011
10	0,002	1111111110	101010101
11	0,002	11111111110	101010100
		$\bar{n}$ = 2,576	$\bar{n}$ = 2,499
		α = 94,54%	α = 97,45%
H(A) = 2,435			1476

Fig. 4 — Esempio di codifica mediante Comma Code e Huffman per una sorgente di simboli a_j emessi con una probabilità $p(a_j)$. In questo esempio le parole si differenziano in base al numero di ripetizione del bit «1», mentre il bit «0» funziona da separatore (bibl. 12). Sono riportati, come confronto i valori del numero medio di bit ($\bar{n}$) e dell'efficienza (α) nel caso dei due codici.

I vantaggi nell'uso di tale codice sono: la possibilità di individuare le parole di codice (sincronizzazione di parola) prima della loro decodifica e la bassa probabilità di propagazione dell'errore di trasmissione, infatti l'errore di trasmissione si propaga solo se è interessato il bit di separazione.

Gli svantaggi principali sono invece: mediamente una minore efficienza rispetto al codice di Huffman e la necessità di usare parole molto lunghe con alfabeti aventi

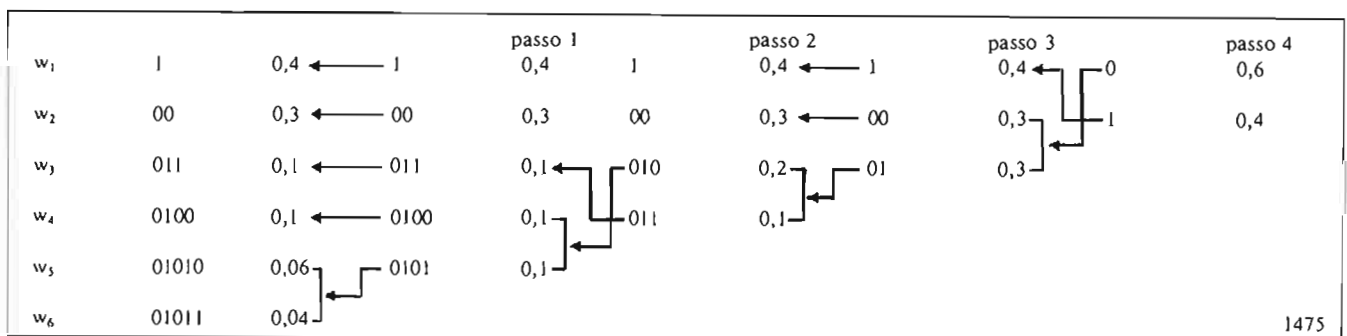



Fig. 3 — Esempio di determinazione delle parole di codice da assegnare ai simboli in entrata. Disponendo della struttura come in figura 2, si procede alla codifica in senso inverso rispetto a quello usato per la costruzione del codice: si inizia ad assegnare «0» ad una delle due ultime probabilità ricavate, normalmente a quella superiore che nell'esempio è 0,6, e «1» all'altra, cioè 0,4; quindi si va un passo indietro decomponendo le probabilità, nell'esempio 0,6 in 0,3 e 0,3, e si aggiunge un bit «0» o «1» alle parole di codice relative alle due probabilità scomposte. La stessa procedura viene ripetuta per ogni passo fino a giungere alle probabilità iniziali dei simboli all'entrata, a cui sono perciò assegnate le parole di codice ricavate e riportate in figura.

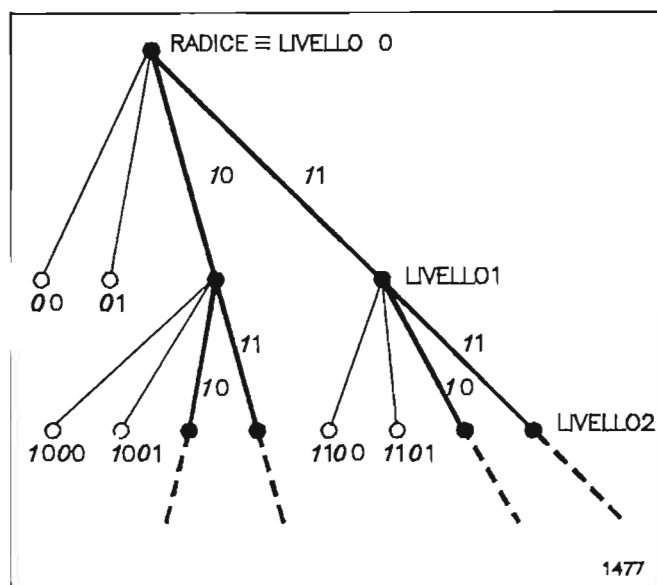



Fig. 5 — Struttura ad albero del B-Code (2,2). In questo caso il blocco elementare è rappresentato da 2 bit, in italico è rappresentato il bit di continuità, ed il numero di diramazioni che proseguono a ciascun livello è anch'esso uguale a 2. Per ciascun livello terminano le diramazioni in cui l'ultima coppia di bit è 00 oppure 01.

un numero elevato di simboli, infatti la parola più lunga ha un numero di bit uguale al numero di simboli dell'alfabeto.

3.4 B-CODE

Come per molti dei VLC noti è comodo descrivere il B-Code tramite una struttura ad albero dove i rami terminali codificano i simboli della sorgente e le diramazioni sono etichettate con bit in modo che, seguendo il percorso dalla radice al ramo terminale, si possa conoscere la codifica binaria di quell'elemento dell'alfabeto di sorgente.

Nel caso di B-Code esiste una famiglia di codici possibili identificati da due parametri (a, d). Il primo parametro rappresenta il numero di bit del blocco elementare, cioè l'unità minima che viene aggiunta per distinguere due parole di codice appartenenti a livelli diversi, mentre il secondo indica il numero di diramazioni usate per il proseguimento dell'albero.

Infatti, mentre il codice di Huffman aggiunge un numero variabile di bit ad ogni diramazione, nel B-Code viene aggiunto un numero di bit pari ad a ed inoltre il numero di diramazioni di proseguimento è d.

Un esempio della struttura ad albero del B-Code (2,2), utilizzato per la codifica in ambito televisivo, è riportato in figura 5.

In base ai due parametri del B-Code si possono fare alcuni calcoli per sapere il numero di parole di codice disponibili ai vari livelli dell'albero.

Anche se il B-code prevede nodi di continuazione ad ogni livello, nel caso in cui si debba codificare un alfabeto finito si utilizzano solo le parole di codice fino al generico livello M. L'albero di codifica termina con tale livello, che può quindi contenere solo nodi terminali; le formule riportate si riferiscono a tale caso.

Dato un generico livello $k \geq 1$, essendo $k = 0$ la radice dell'albero, da esso si diramano d^k sottoalberi e si hanno $2^a - d$ terminazioni per ogni sottoalbero. Il numero di parole codificabili al livello k è quindi:

$$n_k = (2^a - d) \cdot d^{(k-1)}$$

e quindi, se si dispone di M livelli, il numero di parole codificabili a livello M, dove si hanno solo terminazioni, è:

$$n_M = 2^a \cdot d^{(M-1)}$$

Se si utilizzano tutte le terminazioni disponibili a livello M e ai livelli inferiori il numero totale di parole codificabili N_T vale:

$$[8] \qquad N_T = \sum_{k=1}^{M-1} nk + n_M =$$

$$= 2^a \cdot \frac{1 - d^M}{1 - d} - d \cdot \frac{1 - d^{(M-1)}}{1 - d} \qquad \text{per } d \neq 1$$

$$N_T = M \cdot 2^a - (M - 1) \qquad \text{per } d = 1$$

Il numero di bit B_M delle parole di codice del livello M risulta:

$$[9] \qquad B_M = a \cdot M$$

Definito il numero di livelli in base al numero di simboli di sorgente bisogna associare le parole di codice a ciascuno di essi.

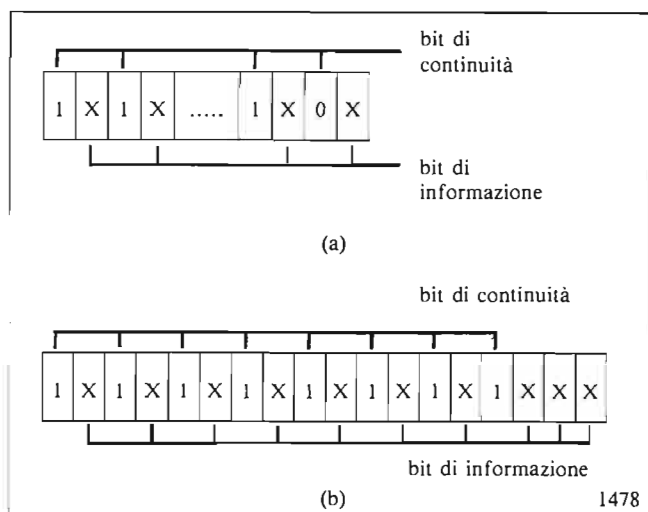



Fig. 6 — Codifica VLC con B2-code. Si distinguono due casi. Le parole di codice con lunghezza compresa tra 2 e 16 bit sono rappresentate in (a). In tale caso è necessario che l'ultimo bit di continuità sia posto a «0» per segnalare la terminazione della parola di codice. In (b) sono rappresentate le parole di codice di lunghezza 18. Tale lunghezza corrisponde a quella massima prefissata per questa applicazione e quindi diventa inutile l'inserimento dell'ultimo bit di continuità che può quindi essere sostituito con un bit d'informazione. Il numero totale di simboli codificabili risulta quindi uguale a $2^{10} + 2^8 \ldots + 2^1 = 1534$.

3.4.1 *Esempio: il B_2 code*

Si considera ora il B-Code (2,2), detto anche *B_2-Code*, nella forma specificata in bibl. 4. Esso si è dimostrato particolarmente adatto alla trasmissione del segnale video codificato mediante DCT.

Tale codice è costituito da blocchi di 2 bit; il primo bit di ogni coppia è detto bit di continuità in quanto indica che la parola continua, bit « 1 », oppure termina, bit « 0 »; il secondo bit di ogni coppia è invece un bit di informazione e permette, al decodificatore, di ricostruire i coefficienti DCT.

In figura 6 è rappresentata la struttura delle parole di codice nel caso specifico del sistema proposto per la codifica del segnale video, in questo caso si è limitata la lunghezza massima delle parole a 18 bit.

Il numero di livelli M dell'albero risulta, in base alla [9], uguale a 9, mentre il numero N_T di parole codificate può essere ottenuto dalla [8] con d = 2 e risulta uguale a 1534.

È inoltre previsto che vengano codificate mediante parole del codice anche le sequenze di zeri ed il simbolo di fine blocco EOB come descritto in 3.1.2.

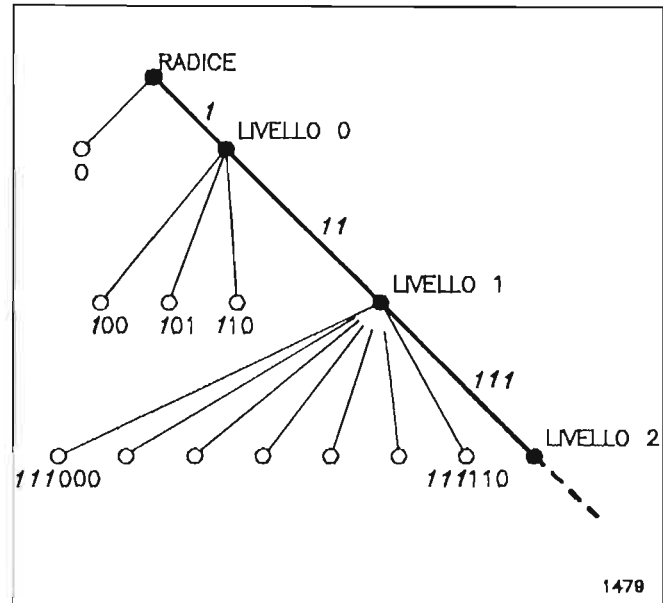



Fig. 7 — Struttura ad albero dell'ACVLC con a = 1 e b = 1. Sono rappresentati tre livelli. I prefissi sono i bit uno tracciati in italico. A livello 2, ad esempio il prefisso è costituito da 11 1 e il suffisso dalle sette combinazioni di tre bit 000-110.

3.5 VLC CALCOLABILE ARITMETICAMENTE

Un altro codice proposto per la codifica del segnale video è l'ACVLC (Arithmetically Computed Variable Length Code).

Questo codice è descrivibile mediante una struttura ad albero in cui esistono rami terminali e nodi di prosecuzione ai vari livelli.

Ad ogni livello k dell'albero vengono aggiunti $\delta l(k)$ bit per codificare i vari simboli:

$$\delta l(k) = a + b \cdot k$$

dove a e b sono due parametri caratteristici dell'ACVLC in esame.

Anche per questo codice si riporta in figura 7 un esempio di struttura ad albero con a = 1 e b = 1.

A livello k una parola di codice è costituita quindi da un numero di bit pari a:

$$l(k) = l(k-1) + \delta l(k)$$

Poiché ad un generico livello k si hanno $n_k = 2^{l(k)} - 1$ rami terminali e 1 nodo di continuazione, codificato con tutti 1, una parola di codice è costituita da k prefissi di lunghezza δl_0, $\delta l_1, \ldots, \delta l_{k-1}$ e un suffisso di lunghezza δl_k; i prefissi sono quindi distinguibili dal suffisso in quanto contengono solo 1. Il numero totale di simboli codificabili con M livelli è:

$$[10] \qquad N_T = \sum_{k=0}^{M} n_k$$

$$= 2^a \cdot \frac{1 - 2^{b(M+1)}}{1 - 2^b} - (M+1) \qquad \text{per } b \neq 0$$

$$N_T = (M+1) \cdot (2^a - 1) \qquad \text{per } b = 0$$

Il numero di bit B_M delle parole di codice del livello M risulta:

$$[11] \quad B_M = \sum_{k=0}^{M} \delta l(k) = a \cdot (M+1) + b \cdot \frac{M \cdot (M+1)}{2}$$

3.5.1. *Esempio: ACVLC proposto per la codifica dei coefficienti DCT*

Si descrive il caso particolare di uso dell'ACVLC come proposto per la codifica del segnale video (bibl. 13).

L'ACVLC viene utilizzato per codificare sia i coefficienti DCT, sia le sequenze di coefficienti nulli.

A partire dalla radice dell'albero il primo bit di codifica serve a distinguere i simboli di sorgente in due categorie: se si tratta di uno 0, il resto della codifica interessa solo sequenze di zeri, se invece è un 1, si tratta di coefficienti diversi da zero.

Nel caso in cui il primo bit sia un 1, cioè si tratta di un coefficiente, il secondo bit indica, se è 0, che il valore è negativo, se è 1, che il valore è positivo. Inoltre il primo bit di codifica di un coefficiente diverso da zero, che segue una parola indicante una sequenza di zeri, viene omesso in quanto superfluo. Non è previsto un simbolo particolare che indichi la fine blocco. Si invia invece separatamente l'informazione del numero di coefficienti codificati per ciascun blocco.

L'ACVLC può essere adattato alla statistica del segnale video variando i parametri caratteristici a e b. In tal caso però, se si varia il codice, bisogna informare il decodificatore dei nuovi parametri usati; il decodificatore deve inoltre conoscere la tabella delle assegnazioni relativa ai nuovi parametri.

Per ridurre la complessità, è stato stabilito che a può variare solo tra 1 e 4 e b tra 0 e 3.

4. Scelta del VLC per la codifica dei coefficienti DCT

4.1 CRITERI DI SCELTA

La scelta del VLC da utilizzare per la codifica del segnale video non è basata su un singolo parametro, ma occorre tenere in considerazione diversi criteri.

Ovviamente il criterio fondamentale è quello dell'efficienza. A questo riguardo il codice ottimo risulta essere quello di Huffman, in quanto minimizza la ridondanza statistica. Sfortunatamente la statistica dell'alfabeto, cioè dei coefficienti DCT, non è perfettamente nota in quanto dipende da vari fattori, per esempio dalla posizione dei coefficienti all'interno del blocco, dal fattore di scalamento, dall'appartenenza alla componente di luminanza o a quella di crominanza, ed il codice di Huffman ha un'efficienza ridotta quando viene utilizzato per codificare alfabeti di sorgente con statistica diversa da quella usata per la generazione del codice.

A questo inconveniente si può ovviare adattando periodicamente il codice alla statistica dei coefficienti, costruendo di conseguenza le parole di codice e riassegnando opportunamente i simboli. Occorre però informare il decodificatore del risultato di queste operazioni e ciò riduce notevolmente sia l'efficienza, sia le prestazioni in caso di errori introdotti dal canale.

Un secondo criterio di scelta è appunto la robustezza del codice nei confronti degli errori. La possibilità di identificare la parola indipendentemente dalla sua decodifica è molto importante a questo riguardo. Un codice tipo «Comma» è ottimale in questo senso, ma offre bassa efficienza di codifica per l'applicazione in esame. Un ultimo criterio è la semplicità realizzativa del decodificatore. Infatti in applicazioni video la velocità di segnalazione è elevata e pertanto è necessario che l'identificazione e la decodifica delle parole di codice avvengano mediante operazioni semplici e possibilmente effettuabili in modo parallelo.

4.2 CONFRONTO FRA I CODICI PROPOSTI

Il sistema descritto brevemente in 3.1.1 è stato proposto per la codifica dei coefficienti DCT. La sua efficienza è simile a quella degli altri candidati, ma in generale sembra meno robusto nel caso di errori, anche se il fatto di consentire livelli di protezione diversi, a seconda del peso del coefficiente DCT codificato, potrebbe essere un vantaggio. Al momento attuale, comunque non sono stati ancora realizzati in pratica i dispositivi di co-decodifica.

Il codice ACVLC descritto in 3.5.1 presenta un'efficienza di poco inferiore a quella del codice di Huffman.

Simulazioni effettuate in Francia, in base alla statistica ottenuta con coefficienti DCT quantizzati in modo non lineare (è stato usato un quantizzatore quasi logaritmico), hanno fornito un valore dell'efficienza pari a circa il 94%.

Si deve considerare che tali risultati sono stati ottenuti utilizzando il codice in modo adattativo. Il codificatore deve quindi attuare delle scelte e comunicarle al decodificatore.

Questo codice è decodificabile più facilmente del codice di Huffman grazie alla divisione delle parole in prefisso e suffisso, ma non consente la sincronizzazione della parola prima della sua decodifica. Infatti l'individuazione della parola di codice si ha solo quando si è giunti ad un nodo terminale dell'albero di decodifica.

Le parole di codice possono inoltre essere facilmente calcolabili, con semplici operazioni aritmetiche, a partire dai parametri determinati dal codificatore e trasmessi al decodificatore.

La lunghezza delle parole di codice varia da 1 a 31 bit.

Un prototipo di codec utilizzante questo codice è stato realizzato e provato recentemente.

Il codice B_2 è stato confrontato con quello ACVLC usando lo stesso alfabeto, proveniente dal quantizzatore logaritmico, di cui si è parlato precedentemente. Si è ottenuta un'efficienza per il B_2-code pari al 91%, inferiore quindi a quella dell'ACVLC.

Il risultato migliore per l'ACVLC è essenzialmente dovuto alla sua adattatività. Infatti, utilizzando quattro differenti B-code, uno del tipo (2,2) e tre del tipo (3,n), e consentendo una scelta adattativa al codificatore, l'efficienza ottenuta è salita a circa il 94%, come nel caso dell'ACVLC.

L'uso adattativo di differenti B-code ha però implicazioni sulla semplicità di decodifica e sulla robustezza nel caso di errori; infatti non essendo la posizione del bit di continuità costante, esistono parole multiple di 2 bit e di 3 bit, le operazioni di sincronizzazione e di decodifica non sono indipendenti.

Dal punto di vista della semplicità di decodifica, il codice B_2 è senz'altro migliore poiché consente l'individuazione delle parole indipendentemente dalla decodifica, limita la propagazione degli errori e consente l'uso di circuiti operanti in modo parallelo. Occorre inoltre notare che la lunghezza delle parole di codice varia da 2 a 18 bit.

Si osservi che nel caso di quantizzatore lineare, come previsto dalle attuali specifiche, l'efficienza calcolata per il B_2-code è di circa il 95% per le sequenze normalmente usate per le simulazioni.

Lo stesso codice, ma con parole di lunghezza massima 16, è utilizzato nel codec di tipo DCT intra-field prodotto e commercializzato dalla Telettra S.p.A. (bibl. 3).

4.3 LA SCELTA

I codici precedentemente descritti sono stati valutati attentamente presso il Centro Ricerche RAI e dalla Telettra S.p.A. in base ai criteri sopra esposti.

Quando si è trattato di decidere il tipo di codice da utilizzare per il sistema di codifica DCT ibrido in via di realizzazione (bibl. 5) la scelta è caduta sul codice B_2.

Si è ritenuto infatti che un codice di tipo adattativo richieda una maggiore complessità del codificatore, che deve essere in grado di operare scelte in tempo reale, e presenti una maggiore sensibilità agli errori introdotti dal canale, occorre infatti fornire informazioni aggiuntive al decodificatore. A fronte di questi svantaggi esiste probabilmente un guadagno in efficienza, la cui entità è comunque di difficile valutazione.

D'altro canto si sono ritenuti fondamentali i criteri di: facile recupero dei sincronismi di parola e di blocco, ro-

bustezza agli errori, semplicità realizzativa. Nei confronti di tali criteri il B_2-code è stato giudicato ineguagliabile.

Per quanto riguarda l'efficienza, il B_2-code presenta un'elevata efficienza, prossima a quella del codice di Huffman.

Si ricorda inoltre che il B_2-code è già stato ampiamente sperimentato con buoni risultati per il codec DCT intra-field.

La scelta del tipo di codice da utilizzare è solo il primo passo per la definizione della struttura del sistema e sua ottimizzazione per la presente applicazione.

5. Ottimizzazione

Il codice prescelto ha la struttura indicata in 3.4.1. In base a simulazioni effettuate mediante calcolatore su sequenze video 4:2:2 è stata determinata l'assegnazione dei valori dei coefficienti DCT e delle sequenze di zeri alle parole di codice.

Dall'analisi statistica dei coefficienti DCT si è osservato che la densità di probabilità dei loro valori ha un andamento di tipo laplaciano ed il valore più probabile è zero; tali risultati confermano studi precedentemente svolti (bibl. 14).

Per ottenere una maggiore efficienza vengono codificate come simboli di sorgente anche le sequenze di zeri. Pertanto i simboli più probabili diventano +1 e −1 a cui si associano le parole lunghe due bit. L'assegnazione procede codificando via via i valori dei coefficienti di valore assoluto più basso e, parallelamente, le sequenze di zeri più corte, che risultano più probabili. Nell'alfabeto di sorgente esistono anche tre simboli particolari indicati con EOB_0, EOB_1 e NULL.

I primi due sono usati per indicare la fine del blocco; nel caso in cui il blocco termini con una sequenza di zeri, questa non viene trasmessa. I due EOB, data l'elevata probabilità di occorrenza, sono codificati con soli 6 bit.

I due simboli vengono usati secondo un ordine prestabilito in modo da consentire il recupero del sincronismo e di rivelare l'eventuale perdita di blocchi (bibl. 15).

Il simbolo NULL viene invece codificato con 12 bit e viene trasmesso per evitare una situazione di svuotamento (underflow) del buffer.

5.1. DETERMINAZIONE DELLE TABELLE DI ASSEGNAZIONE

L'assegnazione delle parole di codice ai valori dei coefficienti ed alle sequenze di zeri è univoca e memorizzata sia nel codificatore che nel decodificatore. Si può evitare di utilizzare una memoria di dimensioni tali da contenere l'assegnazione di tutte le parole di codice. Si è verificato che è sufficiente memorizzare l'assegnazione per i valori inferiori o uguale a 16, in valore assoluto, e per le sequenze di zeri. L'assegnazione ai valori superiori può essere fatta in base ad una legge deterministica che non è strettamente legata alla statistica dei coefficienti; ciò influenza poco le prestazioni del codice, in quanto interessa i valori meno probabili.

Poiché i valori da memorizzare sono limitati si può pensare di utilizzare più tabelle, in funzione del tipo di componente video (luminanza o crominanza), del modo di codifica (intra o inter), del fattore di trasmissione. Si è posto come unico vincolo l'uso delle stesse parole, in tutte le eventuali tabelle, per gli EOB.

Per determinare il numero delle tabelle e i criteri di assegnazione occorre analizzare la distribuzione delle frequenze relative dei valori dei coefficienti DCT, e successivamente determinare le tabelle ottime assegnando le parole di codice più corte ai valori più frequenti.

Il numero N di simboli esaminati ed il numero B di bit per simbolo sono dati da:

$$N = \Sigma f_i$$

$$B = \frac{\Sigma f_i \cdot b_i}{N}$$

dove f_i indica la frequenza assoluta e b_i il numero di bit di ciascun simbolo. Il numero medio B di bit per simbolo è calcolato per le tabelle di assegnazione determinate dall'istogramma delle frequenze dei valori dei coefficienti DCT. Inoltre tale valore è stato determinato assegnando solo le parole con lunghezza fino a 10 bit oppure considerando quelle fino a 12 bit. In quest'ultimo caso sarebbero necessarie, per l'assegnazione delle parole di codice, tabelle di dimensione doppia.

L'analisi statistica effettuata riguarda le parole codificabili relative: al caso in cui non si opera distinzione fra i coefficienti, a quello in cui le due componenti di luminanza e crominanza sono considerate separatamente ed infine al caso in cui sono considerati separatamente i tre modi intra/inter di codifica DCT.

Per poter disporre di risultati significativi, lo studio è stato effettuato su una sequenza di 96 quadri, denominata MIX, che contiene tre spezzoni di sequenze con caratteristiche differenti, per garantire un'informazione indipendente dalle caratteristiche di una singola immagine.

Le simulazioni sono state effettuate a bit-rate costanti, pari a 10 e a 30 Mbit/s; questo affinché la retroazione del buffer garantisca che l'entropia dei coefficienti DCT risulti indipendente dalle immagini trattate.

L'utilizzo delle tabelle ottime di assegnazione in versione ristretta, ossia con parole fino a 10 bit, ed estesa, con parole fino a 12 bit, non hanno comportato variazione significativa dell'efficienza.

La scelta della tabella di assegnazione in funzione del modo intra/inter di trattamento del blocco non fornisce miglioramenti rispetto all'uso di una sola tabella di assegnazione. Ciò implica la non dipendenza della tabella di assegnazione del B_2-code nei confronti dell'uso della compensazione del movimento. Infatti, l'utilizzo di tale tecnica comporta una variazione percentuale dei modi intra/inter in cui sono trattati i singoli blocchi.

Un miglioramento delle prestazioni si ottiene invece utilizzando due diverse tabelle di assegnazione, una per la componente di luminanza e l'altra per quella di crominanza.

In questo caso, l'uso di una tabella distinta per la crominanza, rispetto all'uso di un'unica tabella ottima, comporta una riduzione di circa il 3% relativamente alla sola componente di crominanza a 30 Mbit/s e di circa il 5% a 10 Mbit/s; poiché il numero di simboli della componente

TABELLA 1

ASSEGNAZIONE DELLE PAROLE DI CODICE AI VALORI DEI COEFFICIENTI DI LUMINANZA (Y) E CROMINANZA (C)

<table>
<tr><th>parola di codice</th><th>Y</th><th>C</th><th>parola di codice</th><th>Y</th><th>C</th></tr>
<tr><td>¦0 1¦</td><td colspan="2">1</td><td>¦0 0¦</td><td colspan="2">−1</td></tr>
<tr><td>¦1 1|0 0|</td><td colspan="2">2</td><td>¦1 0|0 1|</td><td colspan="2">−2</td></tr>
<tr><td>¦1 1|0 1|</td><td colspan="2">0*1</td><td>¦1 0|0 0|</td><td colspan="2">0*2</td></tr>
<tr><td>¦1 1|1 0|0 0|</td><td colspan="2">0*3</td><td>¦1 0|1 1|0 1|</td><td colspan="2">0*4</td></tr>
<tr><td>¦1 1|1 0|0 1|</td><td>3</td><td>0*5</td><td>¦1 0|1 1|0 0|</td><td>−3</td><td>0*6</td></tr>
<tr><td>¦1 1|1 1|0 0|</td><td>4</td><td>0*7</td><td>¦1 0|1 0|0 1|</td><td>−4</td><td>0*8</td></tr>
<tr><td>¦1 1|1 1|0 1|</td><td colspan="2">E0B$_1$</td><td>¦1 0|1 0|0 0|</td><td colspan="2">E0B$_0$</td></tr>
<tr><td>¦1 1|1 0|1 0|0 0|</td><td>0*5</td><td>3</td><td>¦1 0|1 1|1 1|0 1|</td><td>0*6</td><td>−3</td></tr>
<tr><td>¦1 1|1 0|1 0|0 1|</td><td>0*7</td><td>4</td><td>¦1 0|1 1|1 1|0 0|</td><td>0*8</td><td>−4</td></tr>
<tr><td>¦1 1|1 0|1 1|0 0|</td><td colspan="2">0*9</td><td>¦1 0|1 1|1 0|0 1|</td><td colspan="2">0*10</td></tr>
<tr><td>¦1 1|1 0|1 1|0 1|</td><td colspan="2">0*11</td><td>¦1 0|1 1|1 0|0 0|</td><td colspan="2">0*12</td></tr>
<tr><td>¦1 1|1 1|1 0|0 0|</td><td colspan="2">5</td><td>¦1 0|1 0|1 1|0 1|</td><td colspan="2">−5</td></tr>
<tr><td>¦1 1|1 1|1 0|0 1|</td><td>6</td><td>0*13</td><td>¦1 0|1 0|1 1|0 0|</td><td>−6</td><td>0*14</td></tr>
<tr><td>¦1 1|1 1|1 1|0 0|</td><td>7</td><td>0*15</td><td>¦1 0|1 0|1 0|0 1|</td><td>−7</td><td>0*16</td></tr>
<tr><td>¦1 1|1 1|1 1|0 1|</td><td>8</td><td>0*17</td><td>¦1 0|1 0|1 0|0 0|</td><td>−8</td><td>0*18</td></tr>
<tr><td>¦1 1|1 0|1 0|1 0|0 0|</td><td>0*13</td><td>6</td><td>¦1 0|1 1|1 1|1 1|0 1|</td><td>0*14</td><td>−6</td></tr>
<tr><td>¦1 1|1 0|1 0|1 0|0 1|</td><td>0*15</td><td>7</td><td>¦1 0|1 1|1 1|1 1|0 0|</td><td>0*16</td><td>−7</td></tr>
<tr><td>¦1 1|1 0|1 0|1 1|0 0|</td><td>0*17</td><td>8</td><td>¦1 0|1 1|1 1|1 0|0 1|</td><td>0*18</td><td>−8</td></tr>
<tr><td>¦1 1|1 0|1 0|1 1|0 1|</td><td colspan="2">0*19</td><td>¦1 0|1 1|1 1|1 0|0 0|</td><td colspan="2">0*20</td></tr>
<tr><td>¦1 1|1 0|1 1|1 0|0 0|</td><td colspan="2">0*21</td><td>¦1 0|1 1|1 0|1 1|0 1|</td><td colspan="2">0*22</td></tr>
<tr><td>¦1 1|1 0|1 1|1 0|0 1|</td><td colspan="2">0*23</td><td>¦1 0|1 1|1 0|1 1|0 0|</td><td colspan="2">0*24</td></tr>
<tr><td>¦1 1|1 0|1 1|1 1|0 0|</td><td colspan="2">0*25</td><td>¦1 0|1 1|1 0|1 0|0 1|</td><td colspan="2">0*26</td></tr>
<tr><td>¦1 1|1 0|1 1|1 1|0 1|</td><td colspan="2">0*27</td><td>¦1 0|1 1|1 0|1 0|0 0|</td><td colspan="2">0*28</td></tr>
<tr><td>¦1 1|1 1|1 0|1 0|0 0|</td><td colspan="2">9</td><td>¦1 0|1 0|1 1|1 1|0 1|</td><td colspan="2">−9</td></tr>
<tr><td>¦1 1|1 1|1 0|1 0|0 1|</td><td colspan="2">10</td><td>¦1 0|1 0|1 1|1 1|0 0|</td><td colspan="2">−10</td></tr>
<tr><td>¦1 1|1 1|1 0|1 1|0 0|</td><td colspan="2">11</td><td>¦1 0|1 0|1 1|1 0|0 1|</td><td colspan="2">−11</td></tr>
<tr><td>¦1 1|1 1|1 0|1 1|0 1|</td><td colspan="2">12</td><td>¦1 0|1 0|1 1|1 0|0 0|</td><td colspan="2">−12</td></tr>
<tr><td>¦1 1|1 1|1 1|1 0|0 0|</td><td colspan="2">13</td><td>¦1 0|1 0|1 0|1 1|0 1|</td><td colspan="2">−13</td></tr>
<tr><td>¦1 1|1 1|1 1|1 0|0 1|</td><td colspan="2">14</td><td>¦1 0|1 0|1 0|1 1|0 0|</td><td colspan="2">−14</td></tr>
<tr><td>¦1 1|1 1|1 1|1 1|0 0|</td><td colspan="2">15</td><td>¦1 0|1 0|1 0|1 0|0 1|</td><td colspan="2">−15</td></tr>
<tr><td>¦1 1|1 1|1 1|1 1|0 1|</td><td colspan="2">16</td><td>¦1 0|1 0|1 0|1 0|0 0|</td><td colspan="2">−16</td></tr>
</table>

Nota: Tabella di assegnazione delle parole B2-code ai simboli più probabili di lunghezza fino a 10 bit. Ai valori inferiori a − 16 e superiori a 16, viene sommato un offset rispettivamente di − 34 e + 33. I bit di informazione nel caso dei valori positivi sono il complemento a quelli dei valori negativi corrispondenti. Analoga situazione si verifica per le sequenze di zeri di ordine dispari e pari e per i due simboli E0B$_0$ e E0B$_1$.
Si precisa che, in questa come nella Tabella successiva, l'asterisco indica una successione di zeri pari al numero che lo segue; ad esempio 0*3 indica 000.

TABELLA 2

ASSEGNAZIONE PER LE SEQUENZE DI ZERI

<table>
<tr><th>PAROLA DI CODICE</th><th>SEQ. ZERI</th><th>PAROLA DI CODICE</th><th>SEQ. ZERI</th></tr>
<tr><td>¦1 1|1 0|1 0|1 0|1 0|0 0|</td><td>0*29</td><td>¦1 0|1 1|1 1|1 1|1 1|0 0|</td><td>0*30</td></tr>
<tr><td>¦1 1|1 0|1 0|1 0|1 0|0 1|</td><td>0*31</td><td>¦1 0|1 1|1 1|1 1|1 1|0 1|</td><td>0*32</td></tr>
<tr><td>¦1 1|1 0|1 0|1 0|1 1|0 0|</td><td>0*33</td><td>¦1 0|1 1|1 1|1 1|1 0|0 0|</td><td>0*34</td></tr>
<tr><td>¦1 1|1 0|1 0|1 0|1 1|0 1|</td><td>0*35</td><td>¦1 0|1 1|1 1|1 1|1 0|0 1|</td><td>0*36</td></tr>
<tr><td>¦1 1|1 0|1 0|1 1|1 0|0 0|</td><td>0*37</td><td>¦1 0|1 1|1 1|1 0|1 1|0 0|</td><td>0*38</td></tr>
<tr><td>¦1 1|1 0|1 0|1 1|1 0|0 1|</td><td>0*39</td><td>¦1 0|1 1|1 1|1 0|1 1|0 1|</td><td>0*40</td></tr>
<tr><td>¦1 1|1 0|1 0|1 1|1 1|0 0|</td><td>0*41</td><td>¦1 0|1 1|1 1|1 0|1 0|0 0|</td><td>0*42</td></tr>
<tr><td>¦1 1|1 0|1 0|1 1|1 1|0 1|</td><td>0*43</td><td>¦1 0|1 1|1 1|1 0|1 0|0 1|</td><td>0*44</td></tr>
<tr><td>¦1 1|1 0|1 1|1 0|1 0|0 0|</td><td>0*45</td><td>¦1 0|1 1|1 0|1 1|1 1|0 0|</td><td>0*46</td></tr>
<tr><td>¦1 1|1 0|1 1|1 0|1 0|0 1|</td><td>0*47</td><td>¦1 0|1 1|1 0|1 1|1 1|0 1|</td><td>0*48</td></tr>
<tr><td>¦1 1|1 0|1 1|1 0|1 1|0 0|</td><td>0*49</td><td>¦1 0|1 1|1 0|1 1|1 0|0 0|</td><td>0*50</td></tr>
<tr><td>¦1 1|1 0|1 1|1 0|1 1|0 1|</td><td>0*51</td><td>¦1 0|1 1|1 0|1 1|1 0|0 1|</td><td>0*52</td></tr>
<tr><td>¦1 1|1 0|1 1|1 1|1 0|0 0|</td><td>0*53</td><td>¦1 0|1 1|1 0|1 0|1 1|0 0|</td><td>0*54</td></tr>
<tr><td>¦1 1|1 0|1 1|1 1|1 0|0 1|</td><td>0*55</td><td>¦1 0|1 1|1 0|1 0|1 1|0 1|</td><td>0*56</td></tr>
<tr><td>¦1 1|1 0|1 1|1 1|1 1|0 0|</td><td>0*57</td><td>¦1 0|1 1|1 0|1 0|1 0|0 0|</td><td>0*58</td></tr>
<tr><td>¦1 1|1 0|1 1|1 1|1 1|0 1|</td><td>0*59</td><td>¦1 0|1 1|1 0|1 0|1 0|0 1|</td><td>0*60</td></tr>
<tr><td>¦1 1|1 1|1 0|1 0|1 0|0 0|</td><td>0*61</td><td>¦1 0|1 0|1 1|1 1|1 1|0 0|</td><td>0*62</td></tr>
<tr><td>¦1 1|1 1|1 0|1 0|1 0|0 1|</td><td>0*63</td><td>¦1 0|1 0|1 1|1 1|1 1|0 1|</td><td>NULL</td></tr>
</table>

di crominanza è decisamente inferiore a quello della luminanza, le riduzioni di bit-rate effettive risultano dello 0,4% e 0,7%.

In base ai risultati delle simulazioni si è pervenuti all'assegnazione delle parole di codice riportate in Tabella 1 e Tabella 2.

L'efficienza riscontrata nelle simulazioni utilizzando tali tabelle è pari a circa il 95%. L'alfabeto è costituito dai valori dei coefficienti DCT ottenuti dal codificatore ibrido ad un bit-rate di 10 e di 30 Mbit/s e quantizzati linearmente.

5.2 Eliminazione del valore + 1 compreso fra sequenze di zeri

Al fine di migliorare l'efficienza di codifica si propone la non trasmissione del valore + 1 quando questo è compreso tra due sequenze di zeri, oppure tra una sequenza di zeri ed il simbolo di fine blocco. Il ricevitore ricostruisce il valore + 1 non trasmesso ogni volta che riceve due sequenze di zeri consecutive oppure una sequenza di zeri seguita da un EOB.

Al fine di valutare la riduzione del numero medio di bit per pel, si è simulato il sistema di DCT Ibrido, sia con la trasmissione che con la non trasmissione del valore + 1 compreso tra sequenze di zeri. La simulazione è stata effettuata utilizzando due differenti sequenze e fattore di trasmissione costante, al fine di eliminare l'influenza del buffer di trasmissione. In questo modo l'unico parametro che varia è il numero medio di bit per pel, mentre rimane costante il rapporto S/N della sequenza trattata, e quindi rimane agevole il confronto delle prestazioni.

In Tabella 3 sono riportati i risultati della simulazione.

La riduzione assoluta di bit per pel aumenta all'aumentare del bit-rate, mentre quella percentuale aumenta riducendo il bit-rate di trasmissione e risulta ≈ 4,5% con un bit-rate equivalente a circa 10 Mbit/s e di ≈ 2,5% con un bit-rate equivalente a circa 35 Mbit/s.

5.3 Conseguenze della scelta del VLC sulla struttura della trama video

La possibilità di utilizzare un codice dalla struttura semplice e robusta come il B_2-code ha consentito di realizzare una organizzazione di trama particolarmente efficiente. L'uso di due parole VLC distinte per il simbolo di fine blocco, cioè EOB_0 e EOB_1, permette di ottenere la sincronizzazione del segnale video anche in presenza di elevati tassi d'errore sul canale, e ciò con un minimo incremento della ridondanza. I pacchetti su cui è basata la struttura di trama sono inoltre progettati anche in funzione delle caratteristiche del codice VLC impiegato (bibl. 15).

6. Conclusioni

Il ruolo della codifica con parole a lunghezza variabile dei coefficienti DCT è fondamentale nel determinare le prestazioni del sistema di riduzione della ridondanza.

La scelta del tipo di codice utilizzato ha implicazioni non solo sull'efficienza di codifica, ma anche sul comportamento del sistema nella fase di acquisizione e mantenimento dei sincronismi e in caso di errori introdotti dal canale.

Il VLC scelto per la realizzazione di codec TV e HDTV e denominato B_2-code presenta ottime caratteristiche sotto questi punti di vista. Consente inoltre la realizzazione di circuiti di decodifica semplici ed estremamente efficienti, anche ad elevate velocità di segnalazione.

Altri elementi del sistema, quali la tecnica di rivelazione degli errori e l'organizzazione della trama video, sono stati progettati tenendo conto delle caratteristiche del VLC.

Fortunatamente il B_2-code presenta, oltre a tutti i vantaggi suddetti, anche un'efficienza di codifica elevata, analoga a quella degli altri codici proposti e prossima a quella di Huffman.

(3847)

TABELLA 3
ELIMINAZIONE DEI VALORI + 1

SEQ	BIT RATE	S/N [dB] Y	C	[Bit/pel] NO	SI	DIFF.	DIFF. %
A	~35	42,5	41,4	3,540	3,455	0,085	2,40
	~20	36,7	38,1	2,076	2,004	0,072	3,46
	~10	31,5	35,7	1,158	1,105	0,053	4,54
B	~35	42,5	41,7	3,480	3,392	0,088	2,51
	~20	36,9	38,1	2,017	1,922	0,068	3,35
	~10	31,9	35,5	1,129	1,082	0,048	4,22

Nota: Numero medio di bit per pel necessario a codificare due sequenze, MIX quadri 1-50 (A) e RENATA BUTTERFLY quadri 2-90 (B), nel caso in cui vengano eliminati SI oppure NO i coefficienti uguali a + 1 compresi fra due sequenze di zeri. I dati sono riportati per tre differenti fattori di trasmissione equivalenti ad un bit-rate medio pari a circa 35, 20 e 10 Mbit/s. Sono indicate le differenze in numero di bit (DIFF) ed in percentuale (%) fra i due casi.

BIBLIOGRAFIA

1. - DEL PERO: *Codifica del segnale televisivo numerico: metodi di riduzione della ridondanza.* «Elettronica e Telecomunicazioni», n. 1, 1989, pagg. 9-16.

2. - BARBERO M., STROPPIANA M.: *Codifica del segnale televisivo numerico: sistemi di riduzione della ridondanza mediante l'uso della trasformata coseno discreta.* «Elettronica e Telecomunicazioni», n. 1, 1989, pagg. 17-29.

3. - *DTV-34 DCT Codec TV per sistemi di trasmissione digitale a 34 Mbit/s.* Telettra - Divisione trasmissione, Febbraio 1989.

4. - IWP CMTT/2: *Revised version of Draft Report AD/CMTT: Digital Transmission of Component-Coded Television Signal at 30-34 Mbit/s and 45 Mbit/s.* Giugno 1989.

5. - BARBERO M., CUCCHI S.: *Codifica del segnale televisivo numerico: architettura di un co-decodificatore utilizzante la DCT.* «Elettronica e Telecomunicazioni», in questo numero.

6. - BENEDETTO S., BIGLIERI E., CASTELLANI V.: *Digital Transmission Theory*, Prentice Hall, 1987.

7. - GRAY R.M.: *Vector Quantization* «IEEE ASSP Magazine», April 1984.

8. - MACQ B., DELOGNE P.: *Progressive Transmission of Pictures by Transform Coding.* «Picture Coding Symposium PCS88», Torino, Settembre 1988.

9. - ROSENFELD A., KAK A.C.: *Digital Picture Processing.*Academic Press.

10. - GONZALES R.C., WINTZ P.: *Digital Image Processing.*Addison Wesley Publishing Company, 2nd Edition.

11. - HUFFMAN D.A.: *A Method for the Construction of Minimum Redundancy Codes.* «Proceedings of the IRE», settembre 1952.

12. - WADE J.G.: *Signal Coding and Processing: an Introduction Based on Video Systems.* Plymouth Politechnic, Ellis Horwood Limited Publischer, Chichester.

13. - TOURTIER P.J., PERRON C.: *Arithmetically Computed Variable Length Codes.* «Picture Coding Symposium PCS88», Torino, Settembre 1988.

14. - REINENGER R.C., GIBSON J.D.: *Distributions of the Two Dimensional DCT Coefficients for Images.* «IEEE Transactions on Communications», vol. COM-31, n. 6, Giugno 1985.

15. - BARBERO M., DEL PERO R., GIROMINI P.: *Codifica del segnale televisivo numerico: struttura di trama per un sistema basato su DCT.* «Elettronica a Telecomunicazioni», in questo numero.

ITALIA '90: UN PASSO SIGNIFICATIVO VERSO LA TELEVISIONE AD ALTA DEFINIZIONE

GIANFRANCO BARBIERI*

SOMMARIO — La televisione ad alta definizione: un servizio che darà agli utenti immagini di straordinaria nitidezza e fedeltà, una rivoluzione tecnologica che coinvolge interessi industriali per migliaia di miliardi. In occasione dei Campionati Mondiali di Calcio (Italia '90), la RAI ha offerto una dimostrazione del grado di spettacolarità e coinvolgimento che questo nuovo mezzo è in grado di fornire; in prima mondiale è stato attuato, su base sperimentale, un collegamento numerico punto-multipunto via satellite che ha consentito di trasmettere in tempo reale le partite a un certo numero di punti di visione dotati di grande schermo. Il sistema di collegamento è stato concepito presso il Centro Ricerche della RAI e sviluppato in collaborazione con l'industria manufatturiera nazionale.

SUMMARY — ***A significant step towards the high definition television:*** *- HDTV: a new service which offers pictures of extraordinary revolution and fidelity, a technological revolution which involves a considerable amount of industrial businesses. During the FIFA World Cup (Italia '90) RAI has given a demonstration of the spectacular performances that this new tool can offer; for the first time in the world, a fully digital chain for point-to-multipoint transmission of HDTV via satellite chain has been put into operation for transmitting in live the football matches to a number of viewing rooms, provided with large screen. The transmission chain has been conceived by the Research Centre of the RAI and developed in co-operation with the national manufacturers.*

La televisione ad alta definizione è argomento di cui si parla con grande interesse ormai da diversi anni, in particolare dall'inizio degli anni '80, da quando cioè i giapponesi cominciarono a dimostrare la fattibilità di vere e proprie produzioni televisive realizzate con questa nuova tecnologia.

Tra le novità emergenti dall'imponente trasformazione oggi in atto nei grandi mezzi di comunicazione di massa, l'HDTV è stata identificata come area di grande rilevanza strategica in quanto essa costituisce un passo significativo verso la TV della «terza generazione». Come già l'avvento del colore ha operato a suo tempo una profonda evoluzione del sistema televisivo, così l'HDTV può rappresentare un evento altrettanto eccezionale: immagini di straordinaria nitidezza, colori più fedeli, visione panoramica su schermi più grandi di quelli convenzionali, il tutto accompagnato da un suono di qualità paragonabile a quello del compact-disc.

La RAI, per tradizione, ha sempre seguito con grande attenzione l'innovazione tecnologica e negli ultimi anni ha ulteriormente accentuato questo suo interesse; l'Alta Definizione, costituisce un modo assolutamente nuovo di produrre che esalta al massimo le potenzialità artistico-espressive del mezzo di ripresa ed incide profondamente sul modo di osservare i programmi.

In questo senso l'impegno della RAI nelle produzioni sperimentali in HDTV rappresenta una precisa scelta a favore del pubblico e della sua domanda di qualità. Per verificare le potenzialità del mezzo, la RAI ha intrapreso fin dall'inizio degli anni '80 una intensa attività sperimentale utilizzando i vari standard di produzione offerti dal mercato, cominciando da quello giapponese e passando a quello europeo (1250 righe/50 semiquadri al secondo) non appena quest'ultimo è stato disponibile.

I Campionati Mondiali di Calcio (*Italia '90*) svoltisi nel giugno scorso hanno offerto un ulteriore severissimo banco di prova per la TV del futuro; due linee di produzione in Alta Definizione sono state attivate e rese operative in modo indipendente: la prima, funzionante a 1250/50, era gestita dalla RAI nell'ambito del progetto europeo EUREKA EU95; la seconda, funzionante a 1125/60, operava in coproduzione tra la RAI e la giapponese NHK.

L'elemento nettamente innovativo della sperimentazione è consistito nell'aver attivato, in anteprima mondiale, un sistema di collegamento punto-multipunto via satellite Olympus interamente numerico, ideato e messo a punto presso il *Centro Ricerche* della RAI. Il sistema, descritto in maggior dettaglio nei vari articoli di questo numero speciale della rivista, era costituito da un complesso di installazioni comprendente tratte in fibra ottica, due stazioni trasmittenti via satellite (una fissa ed una mobile) otto postazioni riceventi attrezzate con proiettori su grande schermo, oltre alle varie unità di monitoraggio e supervisione.

Protagonista dell'esperimento è stato il codificatore HDTV numerico che, operando con sofisticati algoritmi di riduzione di ridondanza, ha reso possibile la trasmissione, senza alterare la qualità delle immagini, attraverso un canale a banda stretta come è quello del satellite Olympus utilizzato per la sperimentazione di RAI-SAT; il codificatore è stato studiato nel quadro del progetto europeo

* Ing. Gianfranco Barbieri del Centro Ricerche RAI (Torino).
Dattiloscritto pervenuto alla Redazione il 23 agosto 1990.

EUREKA EU256 che ha visto come protagonisti la RAI e la Telettra da parte italiana e Retevision, Telettra Espana, Università di Madrid per la parte spagnola. Lo sviluppo in «hardware» dell'apparato è stato curato dalla Telettra.

Un secondo elemento determinante, ai fini della fattibilità della sperimentazione, è stato la ottimizzazione dell'intero collegamento numerico effettuato presso il *Centro Ricerche* RAI tramite simulazione al calcolatore e prove sperimentali. Gli studi suddetti sono stati condotti analizzando le caratteristiche delle varie unità componenti il collegamento. Sulla base dei risultati ottenuti, sono state sviluppate apposite unità riceventi, per ora non disponibili sul mercato, adatte ad operare con segnali numerici; si sono inoltre ottimizzate le prestazioni della stazione fissa (che in origine era stata progettata per trasmissioni TV/MF analogiche) e si sono definiti i dati di progetto della stazione mobile.

L'intera catena di trasmissione è stata progettata, grazie alla estrema flessibilità dei sistemi numerici, per operare in modo trasparente ai differenti formati di scansione attualmente in discussione presso gli organismi internazionali; ciò ha reso possibile la visione degli eventi ripresi da entrambe le linee di produzione.

Nel corso dell'intera manifestazione sono state riprese e trasmesse in diretta 16 partite più un certo numero di repliche in differita; poiché la capienza media delle sale visione si aggirava intorno alle 100 ÷ 150 persone, un totale di circa 1000 persone per partita ha potuto assistere agli esperimenti. Il grado di coinvolgimento degli spettatori era particolarmente evidenziato dalla vistosa partecipazione emotiva alle fasi agonisticamente significative dei vari incontri.

La qualità delle immagini riprodotte, valutata sia sugli schermi grandi che sui monitori professionali di controllo, è sempre risultata paragonabile alla qualità da studio; il funzionamento dell'intera catena di trasmissione numerica si è rivelato pienamente affidabile anche in condizioni di propagazione decisamente critiche a causa di temporali.

Il successo della sperimentazione ha dimostrato la fattibilità tecnico-operativa di un sistema totalmente numerico per il trasporto punto-multipunto, via fibra ottica e satellite, di programmi HDTV con il mantenimento della qualità tecnica della sorgente. La sperimentazione è stata condotta con la partecipazione ed il contributo dell'Istituto Superiore del Ministero delle Poste e delle Telecomunicazioni.

La disponibilità di un tale sistema è suscettibile di avere un impatto non trascurabile nel decollo, su base industriale, delle tecnologie legate alla HDTV.

Nel trattare le problematiche di sviluppo della HDTV si fa generalmente riferimento all'applicazione più ovvia: la diffusione al grande pubblico. Nonostante il grande sforzo prodotto dall'industria "consumer", sia europea che giapponese, si ritiene che siano necessari ancora alcuni anni prima che i ricevitori domestici HDTV vengono resi disponibili sul mercato a prezzi accettabili dall'utenza. Le società di produzione dei programmi devono, tuttavia, programmare adeguati investimenti negli impianti con notevole anticipo rispetto alle date previste per l'attuazione di un servizio rivolto al grande pubblico; la rete dei collegamenti, cioè l'infrastruttura che consente di trasportare il segnale verso impianti di produzione distanti, è solitamente parte integrante del sistema di produzione.

In realtà, lo sviluppo dell'Alta Definizione, sembra avvalersi anche di una diversa prospettiva; i vantaggi da essa offerti in termini di qualità e di processabilità dell'immagine contribuiscono alla promozione di un certo numero di applicazioni sia in campo industriale che in quello commerciale, le quali, essendo meno vincolate da fattori di costo, avranno verosimilmente una penetrazione più rapida. Tra le più importanti, vale la pena citare la telemedicina, la telediagnostica, le applicazioni in campo educativo, la distribuzione di messaggi audiovisivi a grandi gruppi di utenti (sale d'aspetto, ristoranti, hotel, discoteche), la telesorveglianza industriale, le applicazioni militari. Molte di tali applicazioni «non broadcast» sono in grado di sfruttare al massimo le potenzialità offerte dall'HDTV avvalendosi della codifica e trasmissione numerica via fibra ottica o satellite.

L'elenco degli operatori industriali che sono stati coinvolti, in varia misura, nell'allestimento della sperimentazione suggerisce un'ultima importante considerazione. Oltre alla già citata Telettra, partner della RAI nel progetto EU256 e artefice dello sviluppo dei codec, le due stazioni trasmittenti, di proprietà, rispettivamente, Telespazio e RAI, sono state costruite dalla Selenia Spazio; i proiettori HDTV su grande schermo sono stati forniti in gran parte dalla Seleco. Non può essere sottovalutata la rilevanza strategica di un esperimento che ha visto operare in sinergia il settore del servizio pubblico ed alcune tra le più significative realtà dell'industria elettronica nazionale nello sviluppo di una tecnologia destinata ad avere, negli anni novanta, un impatto colossale sul mercato mondiale. (3907)

ITALIA '90: PRIMA MONDIALE DI COLLEGAMENTO NUMERICO IN HDTV VIA SATELLITE

M. ARDITO, G. BARBIERI, M. COMINETTI*

SOMMARIO — I Campionati Mondiali di Calcio disputatisi recentemente in Italia, oltre ad aver rappresentato il più importante evento sportivo del 1990, hanno offerto l'occasione per sperimentare la fattibilità di un sistema di trasmissione punto-multipunto via satellite Olympus interamente numerico. Nel corso della manifestazione, sono state riprese in HDTV e trasmesse «in diretta» 16 partite. I segnali erano generati negli stadi si calcio di varie città con due catene di produzione indipendenti operanti rispettivamente a 1250 linee/50 Hz e 1125 linee/60 Hz e venivano codificati in forma numerica a 70 Mbit/s, impiegando il codec sviluppato nel quadro del progetto europeo Eureka EU-256. Trasmessi con modulazione QPSK tramite il satellite Olympus, erano ricevuti e presentati su grande schermo in sale di visione appositamente attrezzate, otto installate in Italia e due in Spagna. La sperimentazione, concepita, sviluppata e coordinata dal Centro Ricerche della RAI, era basata sulla realizzazione, in prima mondiale, di un sistema di trasporto del segnale HDTV interamente numerico che comprendeva il canale da satellite e collegamenti in fibra ottica. I componenti chiave del sistema, quali il codec HDTV ed i sistemi riceventi, sono stati realizzati in sinegia con le realtà industriali nazionali.
L'articolo descrive la configurazione del sistema di trasmissione, le modalità degli esperimenti nel contesto più generale di Italia '90 e presenta le possibili applicazioni del sistema sia in campo « broadcast» che in campo «non broadcast».

*SUMMARY — **Italia '90: World premiére of a fully digital HDTV broadcasting system via satellite:** - The recent FIFA World Cup 1990 (Italia '90) marked the occasion of experimenting the feasibility of a fully digital system for the point-to multipoint HDTV transmission via Olympus satellite. In this circumstance 16 football matches were shot and live transmitted. The signals were generated in the stadiums of various towns by two independent production chains, operating at 1250 lines/50 Hz and 1125 lines/60 Hz, respectively, and were digitally encoded at 70 Mbit/s by means of the codec developed in the framework of the EUREKA EU-256 European Project. They were transmitted with QPSK modulation via Olympus satellite to eight viewing theatres in Italy and two in Spain, all equipped with large screen projectors. The experiment, conceived, developed and co-ordinated by the RAI Research Centre, was based on the implementation, first in the world, of a fully digital HDTV signal transport system which included the satellite channel as well as an optical fiber link. The main elements of the system, such as the HDTV codec and the receiving parts were implemented in synergy with the Italian industrial realities.*
The paper describes the configuration of the transmission system, the experiment conditions in the more general context of Italia '90 and describes the possibile applications both in «broadcast» and «non broadcast» fields.

1. Introduzione

I satelliti sono stati i grandi protagonisti della diffusione delle immagini dei Campionati Mondiali di calcio e, grazie al loro impiego, i segnali generati in dodici stadi italiani, teatro del più importante evento sportivo del 1990, hanno potuto essere distribuiti in tutto il mondo. Se, per quanto concerne la televisione convenzionale, la fitta rete di collegamenti ha svolto la funzione di soddisfare pressanti esigenze di diffusione di massa a livello mondiale, per quanto concerne la TV ad Alta Definizione, la disponibilità del satellite Olympus ha offerto l'occasione di allestire un complesso sistema di trasmissione, comprendente collegamenti in fibre ottiche, per il trasporto in tempo reale delle immagini HDTV riprese in vari stadi, dal luogo di generazione ad un certo numero di sale attrezzate con grandi schermi e dislocate sul territorio nazionale.

Già in occasione delle Olimpiadi di Seul '88, la NHK aveva organizzato un esperimento similare, seppure di portata più limitata, utilizzando come sistema di trasmissione HDTV il formato MUSE.

Nel caso di Italia '90 occorre tuttavia sottolineare il carattere nettamente innovativo delle dimostrazioni: la sperimentazione, concepita, sviluppata e coordinata dal *Centro Ricerche* RAI, era basata, per la prima volta nel mondo, su un sistema di trasporto del segnale HDTV via satellite interamente numerico; i componenti chiave della catena di trasmissione sono stati progettati e realizzati in sinergia con le realtà industriali nazionali.

* Pi. Maurizio Ardito, ing. Gianfranco Barbieri, dott. Mario Cominetti del Centro Ricerche RAI - Torino.
Dattiloscritto pervenuto alla redazione il 12 settembre 1990.

Lo sviluppo del sistema, nella configurazione impiegata a Italia '90, è stato reso possibile dall'esistenza del codec numerico HDTV in grado di ridurre il flusso binario da circa 1 Gbit/s fino a 70 Mbit/s, senza apprezzabile degradazione del segnale di sorgente; una tale velocità di trasmissione associata ad una modulazione QPSK ha consentito di accedere, previa ottimizzazione dell'intera catena di trasmissione, al canale 24 del satellite Olympus. La disponibilità di un canale da satellite ad elevata potenza e copertura nazionale, ha dimostrato, in definitiva, la fattibilità di un collegamento punto-multipunto di ricezione su tutto il territorio; un ulteriore punto di ricezione è stato allestito in Spagna, a Barcellona, dalla Retevision nonostante la collocazione geografica ampiamente marginale rispetto all'area di copertura dell'Olympus, a riprova della notevole affidabilità del sistema.

Il codec numerico, grazie alla flessibilità della sua architettura interna, è in grado di funzionare in modo trasparente ai vari formati di scansione HDTV oggi in discussione al CCIR (Comitato Consultivo Internazionale di Radiocomunicazioni). Tale prerogativa ha offerto lo spunto per conferire un'ulteriore elemento di originalità alla sperimentazione: il sistema di trasmissione era in grado di operare indifferentemente nei due formati di scansione attualmente in discussione al CCIR, secondo i quali sono state effettuate le riprese (1250 righe/50 Hz e 1125 righe/60 Hz).

Il presente articolo descrive la configurazione del sistema di trasmissione, le modalità degli esperimenti, la loro articolazione nel contesto del più generale evento rappresentato dai Mondiali di Calcio ed inoltre in esso vengono discusse le possibili applicazioni del sistema, sia in campo «broadcast», che in campo «non-broadcast».

2. Configurazione dell'esperimento

Il complesso di impianti sviluppati per attuare il collegamento numerico punto-multipunto era strutturato secondo lo schema di massima di figura 1.

I programmi venivano generati da due catene di produzione indipendenti operanti, rispettivamente, a 1250 righe/50 Hz e 1125 righe/60 Hz.

La prima linea di produzione era gestita dalla RAI nel quadro della sua partecipazione al progetto Eureka EU 95; essa operava in permanenza presso lo stadio Olimpico di Roma e si avvaleva di due pulman di ripresa interconnessi, equipaggiati con apparati funzionanti secondo il formato europeo. Mediante questa linea di produzione si sono effettuate le trasmissioni di sei partite, compresa la finale, giocata a Roma.

Una seconda linea di produzione, installata su una unità mobile della NHK, operava secondo il formato 1125/60 Hz e si spostava per effettuare riprese negli stadi

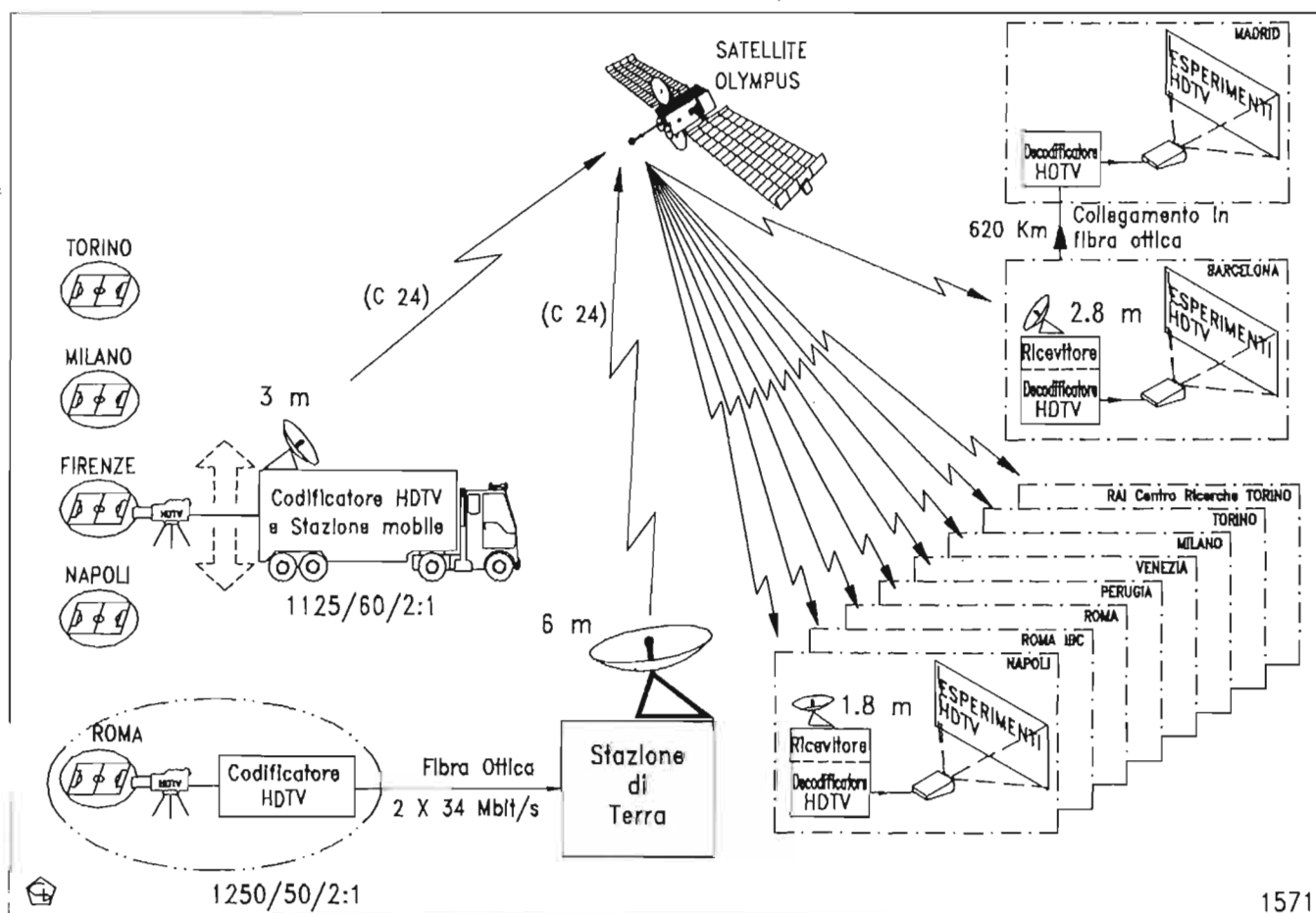



Fig. 1 — Schema riassuntivo degli impianti utilizzati per il collegamento punto-multiplo in HDTV numerica via satellite Olympus.

Fig. 2 — Vista di una delle postazioni di ripresa video (in primo piano la telecamere in HDTV) allo stadio Olimpico di Roma.

di Milano (3 partite trasmesse), Napoli (4 partite inclusa la semifinale), Torino (2 partite), Firenze (1 partita) e Bari (1 partita). Le riprese venivano effettuate in coproduzione RAI-NHK.

La trasmissione verso il satellite Olympus avveniva tramite una stazione fissa, ubicata presso il Centro di Produzione TV di Via Teulada in Roma, per le partite riprese allo Stadio Olimpico e tramite una stazione mobile che si spostava al seguito della unità di ripresa HDTV, per le partite giocate fuori Roma.

I segnali HDTV, generati in bassa frequenza dai mezzi di ripresa, venivano immessi nel sistema di collegamento punto-multipunto attraverso le stazioni trasmittenti attrezzate per il controllo monitorio del segnale proveniente dal satellite; in entrambi i punti di inserzione, il segnale HDTV, proveniente dai programmi di ripresa nel formato in componenti analogiche RGB, veniva convertito in formato numerico e introdotto nel codificatore ove il flusso binario globale, aggirantesi sul Gbit/s, veniva ridotto a 70 Mbit/s; il segnale video numerivo veniva inoltre multiplato ai segnali suono e dati. L'informazione sonora era costituita da tre canali musicali di elevata qualità (CCITT, 384 Kbit/s): una coppia di canali era utilizzata per il suono stereo principale, mentre il terzo canale trasportava le informazioni relative all'effetto ambiente (surround sound).

Nel caso della installazione fissa di Roma, il segnale numerico (audio, video) veniva trasportato dal punto di generazione, presso lo stadio Olimpico, alla stazione trasmittente, installata al Centro di produzione RAI, tramite un collegamento in fibra ottica di lunghezza di circa 2,5 km ed infine inviato al modulatore QPSK per essere trasmesso attraverso il canale 24 del satellite Olympus.

Nel caso della installazione mobile di ripresa, il multiplex erogato dal codificatore HDTV alimentava direttamente il modulatore della stazione trasmittente, operante anch'essa sul canale 24 del satellite Olympus.

Entrambe le stazioni trasmittenti erano attrezzate, a scopi monitori, con un impianto ricevente completo, comprendente il demodulatore QPSK, il decodificatore HDTV digitale, il convertitore digitale/analogico e un monitore di qualità. Dalla sala monitoria del Centro di produzione di Roma il segnale ricevuto dal satellite veniva, inoltre, rinviato tramite collegamento in fibra ottica al punto di inserzione dello stadio Olimpico ove poteva essere visualizzato, a scopi di controllo, dal personale ivi operante.

I segnali numerici HDTV ritrasmessi dal satellite venivano ricevuti in sette postazioni allestite nelle città di Milano, Napoli, Perugia, Roma (2 postazioni), Torino, Venezia. Presso il *Centro Ricerche* RAI a Torino, che funzionava da centrale operativa per coordinare gli esperimenti di trasmissione, era stato inoltre allestito un ulteriore impianto ricevente equipaggiato con vari livelli di monitorie.

La ricezione avveniva, in ciascuna postazione, tramite un'antenna parabolica da 1,8 m di diametro[1], una unità ricevente, il demodulatore QPSK e il decoder HDTV. Il segnale numerico uscente dal decodificatore veniva infine convertito nel formato analogico in componenti RGB per essere distribuito ai proiettori su grande schermo (in alcune postazioni le sale di proiezione erano più di una) e ai vari apparati di controllo.

3. Installazione ricevente allestita in Spagna

Nonostante la Spagna sia in posizione geografica alquanto marginale rispetto all'area di copertura del satellite Olympus (canale 24), è stato ugualmente allestito un punto ricevente a Barcellona a cura di Retevision e Telettra Española, partner di RAI e Telettra nel progetto Eureka EU 256. La postazione era equipaggiata con un'antenna parabolica da 2,8 metri di diametro e con le stesse attrezzature utilizzate in Italia. Il segnale decodificato veniva visualizzato su grande schermo a Barcellona in una sala della capienza di circa 120 persone; il segnale numerico a 70 Mbit/s veniva inoltre inviato, tramite collegamento in fibra ottica di circa 620 km, a Madrid ove era stata allestita una ulteriore sala visione, anch'essa attrezzata con grande schermo. Le proiezioni venivano effettuate a Madrid in concomitanza con la conferenza internazionale Eurotelecom 90 e Expotronica Simposia.

4. Sala visione

Come si è detto, in Italia sono stati attrezzati otto punti riceventi; a Milano, Venezia e Napoli la sala visione era ospitata nelle rispettive sedi RAI; a Roma vi erano due sale, una presso il Centro di produzione Radiofonica della RAI e una presso i nuovi insediamenti RAI di Saxa Rubra (ove durante i Mondiali di Calcio '90 era ospitato l'International Broadcasting Centre), a Perugia la sede della proiezione era situata presso la locale Camera di Commercio ed a Torino presso gli ex stabilimenti industriali della Fiat Lingotto.

Le proiezioni avvenivano, su schermi piani di 3,5 metri di larghezza e 2 di altezza, tramite videoproiettori funzionanti a proiezione diretta mediante tre tubi di tipo CRT. L'illuminazione della sala era ridotta al minimo consentito dalle norme di sicurezza. I videoproiettori erano stati adattati per operare in entrambi i formati di scansione (1250/50 e 1125/60).

Per la disposizione del pubblico si è cercato il più possibile di seguire lo schema di figura 3. Le sale erano anche dotate di un proiettore a 625 linee 50 Hz in grado di proiettare sullo stesso schermo immagini PAL ricevute da canali TV terrestri nel caso di mal funzionamento del collegamento numerico via satellite.

Nonostante l'esperimento sia stato condotto con apparati allo stadio di sviluppo praticamente prototipale, merita segnalare che per tutta la durata dei campionati mondiali non si è mai avuto alcun inconveniente di rilievo e non è stata pertanto mai utilizzata la riserva PAL, pur operando in alcuni casi in condizioni ambientali molto critiche: si sono superati ad esempio i 40 gradi centigradi all'interno della stazione mobile durante una trasmissione da Firenze e si è affettuata una trasmissione da Milano durante un fortissimo temporale con grandine.

L'esperienza acquisita ha confermato che, non potendo utilizzare schermi con forte guadagno, che comportano una illuminazione dello schermo non uniforme e la riduzione dell'angolo di visione utile a causa della direttività che ne consegue, la luminosità ottenibile con proiettori a tre tubi su schermi di 7 metri quadrati, in condizioni di definizione ancora accettabili, non è ancora ottimale per manifestazioni in presenza di pubblico; d'altra parte, se si aumenta la luminosità si perde decisamente in definizione.

Con schermi di 2 metri di altezza, considerando utile al fine di una visione HDTV una distanza da 3 a 7 volte l'altezza dello schermo, gli spettatori in condizioni di visione ottimale sono circa un centinaio. Tenendo conto del numero di partite trasmesse, alcune delle quali replicate in giorni successivi, si può stimare in diecimila il numero complessivo di spettatori che hanno assistito alla manifestazione.

Per quanto concerne l'audio, le sale erano attrezzate con diffusori ai lati dello schermo che fornivano l'audio stereofonico tradizionale, contenente il commento ed una parte di effetti, e una serie di altoparlanti, distribuiti in maniera semicircolare nella parte retrostante della sala, che diffondevano il canale «surround» che contribuiva a completare l'effetto presenza della trasmissione in HDTV su grande schermo.

5. Prestazioni del sistema

L'esperienza acquisita durante la sperimentazione in occasione di Italia '90 dimostra che il pubblico è molto interessato alla visione in HDTV su grande schermo di avvenimenti sportivi tanto da immedesimarsi completamente nell'evento. A ciò ha sicuramente contribuito l'impiego del formato di visualizzazione «allargato» (16:9 contro 4:3 della TV convenzionale). Come conseguenza, spesso il tifo in sala era paragonabile a quello dello stadio di calcio.

Nonostante i limiti ricordati dei proiettori, la qualità dell'immagine co-decodificata è risultata ottima, secondo le aspettative, e corrispondente perfettamente alla qualità stimata attraverso simulazioni mediante calcolatore ed ai risultati di prove soggettive. Infatti l'unico modo per valutare le prestazioni dei codec è quello di ricorrere a prove soggettive, utilizzando speciali sequenze di prova, essendo le degradazioni legate al contenuto dell'immagine e quindi non misurabili con i sistemi tradizionali. Attualmente è molto difficile effettuare prove soggettive con segnali ad alta definizione a causa dei limiti tecnologici delle telecamere e dei display, che non consentono ancora una sufficiente risoluzione, per evidenziare gli eventuali difet-

(1): È stato effettuato anche un esperimento di ricezione al *Centro Ricerche* RAI, utilizzando un'antenna da 70 cm. L'esperimento, avvenuto a cielo sereno, non ha evidenziato alcuna degradazione nella qualità ed affidabilità del segnale ricevuto.

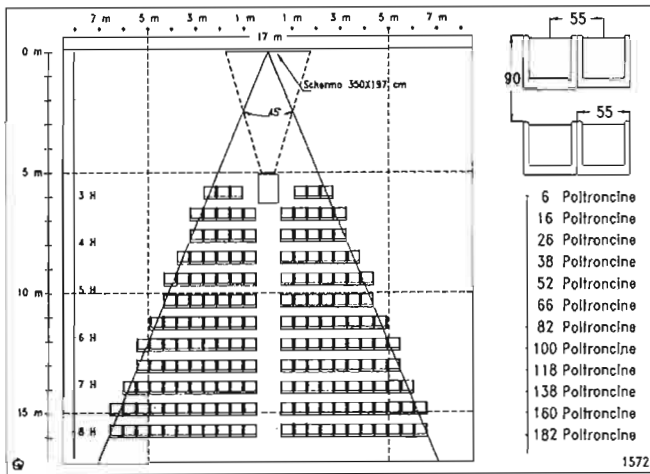



Fig. 3 — Schema della disposizione del pubblico a cui si è fatto riferimento per l'allestimento delle sale visione.

Per quanto concerne le prestazioni del canale via satellite per la trasmissione del segnale numerico HDTV si è riscontrata una sostanziale corrispondenza con i risultati ottenuti, durante la fase preparatoria, tramite accurate simulazioni al calcolatore e misure di laboratorio su un simulatore «hardware» del canale Olympus appositamente realizzato presso il *Centro Ricerche* RAI.

L'unica causa di degradazione, rispetto alla trasmissione su canale ideale, era imputabile, come previsto, alla limitata larghezza di banda del canale Olympus (circa 42 MHz). Tale degradazione, quantizzabile in termini di peggioramento del rapporto portante-rumore (C/N), per un assegnato tasso d'errore, è risultata di circa 2 dB, in un buon accordo con le simulazioni e le prove sperimentali. Ciò non ha comunque sostanzialmente penalizzato i margini di corretta ricezione del segnale, risultati normalmente superiori a 10 dB.

Un'analisi più dettagliata delle prestazioni del collegamento via satellite è riportata in bibliografia 1.

ti degli algoritmi previsti per trasferire un segnale della qualità HDTV. Per risolvere questo problema si effettuano le prove nello standard per la TV digitale convenzionale 4:2:2 considerando l'immagine come una «finestra» su un display più grande in HDTV. Poiché il display virtuale in HDTV avrebbe un'altezza doppia del display reale in 4:2:2, effettuando le prove ad una distanza di sei volte l'altezza dello schermo si ottengono risultati equivalenti ad un'osservazione in HDTV ad una distanza di tre volte l'altezza, il che corrisponde esattamente alle procedure raccomandate dal CCIR per la valutazione delle immagini in HDTV. In figura 4 sono riportati i risultati di prove soggettive effettuate presso il *Centro Ricerche*RAI che, come si è detto, corrispondono perfettamente a quanto osservato durante l'esperimento Italia '90.

6. Applicazioni del sistema di trasmissione

Il codec HDTV impiegato negli esperimenti di Italia '90 è caratterizzato dall'elevato grado di flessibilità che costituisce una delle prerogative dei sistemi numerici e pertanto si adatta a vari formati di scansione e può operare con una gamma relativamente vasta di velocità di trasmissione. Ciò rende possibile l'utilizzo del sistema in un certo numero di applicazioni nel campo dei collegamenti via satellite (bibl. 2):

— scambio di materiali HDTV di elevata qualità tra Centri di Produzione;
— trasmissione da unità di ripresa esterna verso i Centri di Produzione (satellite news gahtering);

Fig. 4 — Risultati di un'indagine soggettiva per la valutazione dei segnali codificati a 34 Mbit/s e 17 Mbit/s, confrontati con l'originale 4:2:2. Si è utilizzato il metodo CCIR doppio stimolo. Le differenze sono state calcolate su base centesimale. Gli osservatori erano disposti ad una distanza pari a 6 volte l'altezza dello schermo, corrispondente a 3 volte, se riferita al caso HDTV.

— trasmissione in diretta di avvenimenti di attualità a utenti selezionati;
— trasmissione di programmi HDTV verso le stazioni di testa di reti di distribuzione in fibra ottica.

Per ciascuna delle suddette applicazioni si richiede uno studio accurato dei requisiti specifici; l'esperienza acquisita durante gli esperimenti di Italia '90 costituisce una base per successive ottimizzazioni del sistema di trasmissione.

L'introduzione di un servizio regolare di diffusione diretta da satellite dell'HDTV all'utente, avverrà secondo modi e tempi dettati principalmente da fattori di mercato legati alla disponibilità dei ricevitori domestici. Per impiego sui canali da satellite a «banda stretta» nella banda di 12 GHz pianificata dalla WARC '77, è allo studio in Europa il sistema HD-MAC. Per quanto riguarda l'eventuale assegnazione per l'HDTV da parte della WARC '92 di una nuova banda, intorno ai 20 GHz, da pianificare con canali a «banda larga» si prevede l'utilizzazione di sistemi completamente numerici, in grado di fornire all'utente una qualità potenzialmente trasparente a quella di studio. I vantaggi offerti dall'HDTV in termini, sia di qualità intrinseca che di processabilità, stanno promuovendo una vasta gamma di interessanti applicazioni di tipo «non broadcast» che potranno presumibilmente essere introdotte prima dell'avvio di un regolare servizio di radiodiffusione all'utente domestico.

Nella maggior parte dei casi, la disponibilità di uno strumento che consenta il trasporto a distanza dei segnali è un requisito determinante per la realizzazione del sistema; tra le applicazioni più significative dell'HDTV in settori «non broadcast» si possono individuare le seguenti:
— applicazioni in campo didattico (insegnamento da postazioni remote);
— pubblicità (programmi di tipo commerciale, cataloghi audiovisivi, sale visioni, teleacquisto);
— applicazioni biomedicali (telemedicina, telediagnostica);
— applicazioni militari;
— videoconferenza ad elevata qualità;
— distribuzione elettronica di film nelle sale cinematografiche;
— punti di visione nei servizi pubblici (stazioni ferroviarie, aeroporti, mostre, conferenze);
— telesorveglianza per usi industriali;
— grafica elettronica e cartellonistica.

Alcune tra le applicazioni sopraelencate sono già in avanzato stato di sviluppo (bibl. 3) non essendo, generalmente, sottoposte a vincoli derivanti dal perdurare del dibattito internazionale sugli standard di scansione per l'HDTV. Se si considera il grado di maturazione tecnologica degli attuali apparati di ripresa e degli attuali display, il sistema di trasmissione dimostrato a Italia '90 rappresenta una valida e versatile risposta alle esigenze di trasporto dei segnali HDTV derivante dall'introduzione dei vari servizi.

7. Conclusioni

Gli esperimenti effettuati dalla RAI durante i campionati mondiali di calcio (Italia '90) hanno dimostrato la fattibilità tecnica e operativa di un sistema di trasmissione totalmente numerico di programmi HDTV di elevata qualità. Le installazioni sviluppate per l'occasione comprendevano tratte in fibra ottica, 2 stazioni trasmittenti via satellite di cui una fissa e una mobile, un collegamento via satellite e un certo numero di piccoli impianti di distribuzione, nelle postazioni riceventi, per alimentare le sale visione equipaggiate con grandi schermi.

Durante gli oltre 40 giorni della manifestazione sono state trasmesse 16 partite in diretta HDTV più alcune repliche; circa 700 persone distribuite in otto sale di visione, distribuite sul territorio nazionale, hanno potuto assistere a ciascuna trasmissione.

Gli esperimenti di trasmissione numerica sono stati resi possibili grazie alla disponibilità del codificatore HDTV sviluppato nel quadro del progetto Eureka EU 256.

In base ai risultati degli esperimenti si possono trarre le seguenti conclusioni:
— la qualità globale delle immagini ricevute, valutata sui monitori HDTV di caratteristiche professionali, è stata giudicata paragonabile a quella esistente in uscita dalle telecamere;
— le prestazioni della catena di trasmissione, in termini di robustezza nei confronti degli errori di trasmissione e di margini di ricezione, sono risultate superiori alle aspettative, anche in condizione di propagazione alquanto sfavorevoli.

Il sistema di trasmissione sperimentato ad Italia '90 si è dimostrato uno strumento adeguato ed efficace per il trasporto a lunga distanza di segnali HDTV di elevata qualità.

Il risultati ottenuti costituiscono un significativo passo avanti nello sviluppo della HDTV. Infatti, la trasmissione punto-punto rappresenta una delle strozzature che occorre eliminare per rendere totalmente operativo un sistema di produzione. Inoltre, la disponibilità di un sistema di trasmissione totalmente numerico, quindi idoneo ad essere impiegato sia su reti in fibra ottica, che su canali da satellite, apre la strada allo sviluppo di una molteplicità di applicazioni industriali e commerciali, nelle quali possono venire sfruttate pienamente la peculiarità intrinseche dell'Alta Definizione.

(3927)

BIBLIOGRAFIA

1. - M. Cominetti, A. Morello: *Italia '90: Trasmissioni punto-multipunto di HDTV via satellite.* «Elettronica e Telecomunicazioni», in questo numero.

2. - M. Cominetti: *Perspectives and evolution of HDTV by satellite.* Workshop on HDTV, Torino, September 1989.

3. - G.F. Barbieri, M. Cominetti: *Experiments of point-to-multipoint digital transmission of HDTV via satellite during the football Worldcup 1990 in Italy.* 132° SMPTE Technical Conference, New York (USA), October 1990.

4. - M. Barbero, S. Cucchi, R. Del Pero, G. Dimino, M. Occhiena, M. Muratori, M. Stroppiana: *Italia '90: Codifica del segnale televisivo numerico.* «Elettronica e Telecomunicazioni», in questo numero.

ITALIA '90: TRASMISSIONE NUMERICA PUNTO-MULTIPUNTO VIA SATELLITE DI SEGNALI HDTV

M. COMINETTI, A. MORELLO*

SOMMARIO — Si esaminano le problematiche tecniche relative alla sperimentazione di trasmissione numerica punto-multipunto via satellite di segnali HDTV a 70 Mbit/s effettuata dalla RAI in occasione della Coppa del Mondo di calcio 1990 in Italia. Particolare attenzione è rivolta alla intensa attività svolta dal Centro Ricerche nella fase preparatoria della sperimentazione per identificare i parametri tecnici del sistema: bit-rate, tipo di modulazione, scelta del satellite, caratteristiche delle stazioni trasmittenti e delle installazioni riceventi. Si riportano quindi i risultati conseguiti nella sperimentazione tecnica ad Italia '90 che hanno sostanzialmente dimostrato l'affidabilità e le elevate prestazioni del sistema a 70 Mbit/s per il trasporto a lunga distanza di segnali HDTV di elevata qualità su canali da satellite e collegamenti in fibra ottica. Si indicano le possibilità di miglioramento delle prestazioni del sistema introducendo la compensazione del movimento nei codec HDTV ed impiegando sistemi avanzati di modulazione e codifica di canale con decodifica di Viterbi, attualmente in fase di sviluppo. I risultati conseguiti nella sperimentazione in Italia '90 consentiranno alla RAI di contribuire fattivamente all'attività di studio e normalizzazione nel campo dell'HDTV da satellite, sia in ambito nazionale (attraverso il Progetto Finalizzato Telecomunicazioni del CNR) che internazionale (CCIR, UER, EU-95 fase II).

SUMMARY — ***Point-to-multipoint digital transmission via satellite of HDTV signals*** *- The technical problems relevant to the experiments of point-to-multipoint digital transmission at 70 Mbit/s via satellite of HDTV signals carried-out by the RAI a the Foot-ball WorldCup in Italy (Italia '90) are discussed. Particular attention is paid to the activity of the RAI Research Centechnical parameters of the digital HDTV system: bit-rate, modulation, choice of the satellite, characteristics of the transmitting earth stations and for the receiving installations. The paper then reports on the results of the experiments at Italia '90 which have shown the high reliability and performance of the system for the long distance transport at 70 Mbit/s of high quality HDTV material over digital satellite and optical fibre links. The possibility of further system improvement, by introducing motion compensation in the HDTV codec and by using advanced channel coding-modulation techniques based on Viterbi decoding, are prsented. The significant results achieved by the Italia '90 experiments will allow the RAI to positively contribute in the development of satellite HDTV, both in the national and internazional contexts (CNR/PFT, CCIR, EBU, EU-95/PG04, etc.).*

1. Introduzione

La Coppa del Mondo di calcio giocata in Italia (8 giugno - 8 luglio 1990) ha dato alla RAI l'opportunità di sperimentare, in anteprima mondiale, un sistema di trasmissione numerica punto-multipunto via satellite di segnali HDTV ad elevata qualità.

La sperimentazione è stata concepita, ideata e coordinata dal *Centro Ricerche* RAI. Lo studio sistemistico è iniziato nel 1988 sotto la guida dell'allora Direttore ing. Rolando Salvadorini, quando fu possibile dimostrare in simulazione, a coronamento di una importante attività di ricerca nel campo della codifica di sorgente, l'elevata qualità delle immagini HDTV ottenibili, con 50-60 Mbit/s assegnati al video, mediante algoritmo di riduzione di ridondanza basato sulla trasformata coseno discreta (DCT). Tale algoritmo di codifica, con l'aggiunta di 2 Mbit/s per la trasmissione dell'audio (5 o 6 canali di qualità) e circa 8 Mbit/s per la correzione degli errori e le informazioni di sincronismo di trama, consentiva di sviluppare un sistema di HDTV numerico con bit-rate lordo di 60-70 Mbit/s, potenzialmente in grado di essere utilizzato per trasmissioni sperimentali su canali da satellite.

La disponibilità dei codec HDTV a 70 Mbit/s, sviluppati in ambito al progetto europeo EU-256 e costruiti dalla Telettra in una decina di esemplari, e la possibilità di accesso diretto al canale 24 sul satellite Olympus da parte della RAI hanno costituito gli elementi chiave del successo della sperimentazione.

Si descrivono nel seguito le varie fasi della complessa attività di ricerca e sperimentazione effettuata dal *Centro Ricerche* RAI in vista di Italia '90, con particolare attenzione alla valutazione delle prestazioni del collegamento numerico via satellite effettuata tramite simulazioni al calcolatore, prove su un simulatore hardware del satellite, appositamente sviluppato presso il *Centro Ricerche*, e con prove preliminari di trasmissione via Olympus.

Si commentano infine i soddisfacenti risultati conseguiti ad Italia '90 esaminando separatamente le prestazioni delle varie unità componenti l'intera catena di trasmissione.

* Dr. Mario Cominetti, Ing. Alberto Morello del Centro Ricerche RAI (Torino).
Dattiloscritto pervenuto alla redazione il 30 ottobre 1990.

2. La fase preparatoria

I primi studi sulle problematiche di trasmissione, in vista della dimostrazione di HDTV numerica ad Italia '90, miravano ad identificare le caratteristiche del canale trasmissivo per satelliti di diffusione diretta, quali l'Olympus, e per satelliti di comunicazione, quali Eutelsat I, e a sviluppare un modello di simulazione al calcolatore (bibl. 1). Si effettuavano quindi le valutazioni per individuare il massimo bit-rate trasmissibile con modulazioni convenzionali tipo QPSK, Offset-QPSK e 8-PSK, per individuare i livelli di interferenza sui canali adiacenti ed i diametri di antenna richiesti per una corretta ricezione del segnale numerico, con adeguati margini anche in condizioni di propagazione sfavorevoli. Per verificare su base sperimentale i risultati ottenuti al calcolatore veniva realizzato un simulatore hardware del canale da satellite, equipaggiato con filtri di entrata (IMUX) ed uscita (OMUX) riproducenti sia quelli dell'Olympus che di Eutelsat I, e con amplificatore di potenza (TWTA) da 50 watt operante nelle bande DBS e KU. Nell'aprile 1989 venivano provati sul simulatore i prototipi a 70 Mbit/s di un modem QPSK e di un modem 8-PSK. Parallelamente veniva avviata da parte della Selenia Spazio (su specifiche definite dal *Centro Ricerche*) la costruzione di una stazione trasmittente trasportabile per il satellite Olympus, con caratteristiche di banda passante e di linearità adatte alle trasmissioni numeriche. Si provvedeva inoltre all'adattamento al numerico della stazione fissa trasmittente installata presso il Centro di produzione RAI-TV di Roma, originariamente progettata per operare esclusivamente con segnali televisivi analogici MF.

Nell'estate del 1989 veniva lanciato il satellite Olympus e nel novembre dello stesso anno venivano effettuate con successo le prime prove di trasmissione numerica via Olympus con segnali pseudo-casuali a 70 Mbit/s in QPSK dalla stazione fissa di Roma. I risultati di queste prove confermavano largamente quelli ottenuti con simulazioni al calcolatore e con prove di laboratorio.

Sulla base dei risultati di bilanci di collegamento l'ipotesi di utilizzare il satellite Eutelsat-I per le trasmissioni HDTV a 70 Mbit/s non sembrava proponibile a causa della bassa potenza irradiata dal satellite che avrebbe richiesto di operare con antenne riceventi molto grandi (in certi casi con puntamento automatico), specialmente nelle intallazioni a sud di Roma, e comunque con scarsi margini di ricezione (vedi par. 3.5). Si sarebbe inoltre dovuto provvedere all'affitto presso Eutelsat di uno o due ripetitori per tutto il periodo della sperimentazione, esigenza particolarmente difficile da soddisfare data la scarsa disponibilità di ripetitori da satellite ad Italia '90. L'ipotesi di utilizzare l'Olympus prendeva quindi via via più consistenza.

Nella primavera del 1990 veniva presentato all'UER a Torino ed al NAB '90 ad Atlanta (USA) il primo prototipo del codec HDTV a 70 Mbit/s sviluppato dalla Telettra nell'ambito del progetto europeo EU-256. La disponibilità di una decina di esemplari del codec, appositamente costruiti dalla Telettra, ha quindi consentito al *Centro Ricerche* RAI di completare entro i termini previsti l'allestimento della intera catena di trasmissione e distribuzione del segnale via satellite e realizzare la sperimentazione di HDTV ad Italia '90.

3. Prestazioni del collegamento via satellite

Un elemento determinante nel successo delle dimostrazioni di HDTV ad Italia '90 è stata certamente l'elevata qualità del collegamento numerico via satellite Olympus ottenuta tramite una rigorosa ottimizzazione delle prestazioni delle singole componenti la catena. In particolare si sono studiati i vari aspetti tecnici che maggiormente potevano influire sulle prestazioni del collegamento:

- tecniche di correzione di errore;
- canale trasmissivo via satellite;
- modulazione numerica;
- problemi interferenziali;
- influenza della propagazione;
- stazioni trasmittenti di terra (fissa e trasportabile);
- sistemi riceventi.

3.1 Correzione degli errori di trasmissione

I codec HDTV impiegati ad Italia '90 utilizzano un codice correttore di tipo BCH(255,239) interlacciato, in grado di correggere errori sigoli e doppi su un blocco di 255 bit. L'interlacciamento (di ordine 16) permette la correzione di errori a pacchetti, fino a un massimo di 32 errori consecutivi. Con tale sistema di correzione l'immagine HDTV non presenta degradamenti fino a un tasso di errore (BER) sul canale di 10^{-6}, a cui corrisponde un BER di circa 10^{-10} dopo correzione. Un BER di 10^{-5} produce un degradamento leggermente disturbante. Il valore di 10^{-6} è stato quindi scelto come riferimento nel dimensionamento dell'intero sistema di trasmissione. Altri codici a blocco, più sofisticati, avrebbero potuto garantire prestazioni migliori: ad esempio il codice di Reed-Solomon (255,239), attualmente in avanzata fase di implementazione nei codec HDTV della seconda generazione, permette un guadagno di oltre 1 dB rispetto al codice BCH nel caso di errori statisticamente indipendenti e consente inoltre di correggere più efficacemente gli errori a pacchetti (fino ad un massimo di 128 bit errati consecutivi).

3.2 Il canale trasmissivo

In figura 1 è riportato lo schema a blocchi della catena di trasmissione HDTV via satellite assunta come riferimento nella fase di studio. Nella stazione di terra il segnale HDTV in uscita dal codificatore su due flussi da 34,368 Mbit/s, viene modulato (in QPSK od 8-PSK) alla frequenza intermedia di 140 MHz, convertito alla frequenza della tratta in salita (rispettivamente a 14 GHz e a 18 GHz per i satelliti Eutelsat-I e per Olympus), amplificato e trasmesso verso il satellite. Per ridurre le interferenze sulla tratta in salita e le distorsioni del segnale, l'amplificatore di potenza (HPA a Klystron) della stazione di terra deve operare a potenza ridotta (almeno 3 dB al di sotto della potenza massima) in regime cioè di quasi-linearità. In queste condizioni si dice che l'HPA opera con un OBO (Output Back Off) di 3 dB.

Il ripetitore del satellite include un ricevitore a basso rumore (LNA), un convertitore di frequenza, un filtro di canale di ingresso (IMUX) necessario per ridurre il rumore sulla tratta in salita, un amplificatore di potenza ad onde

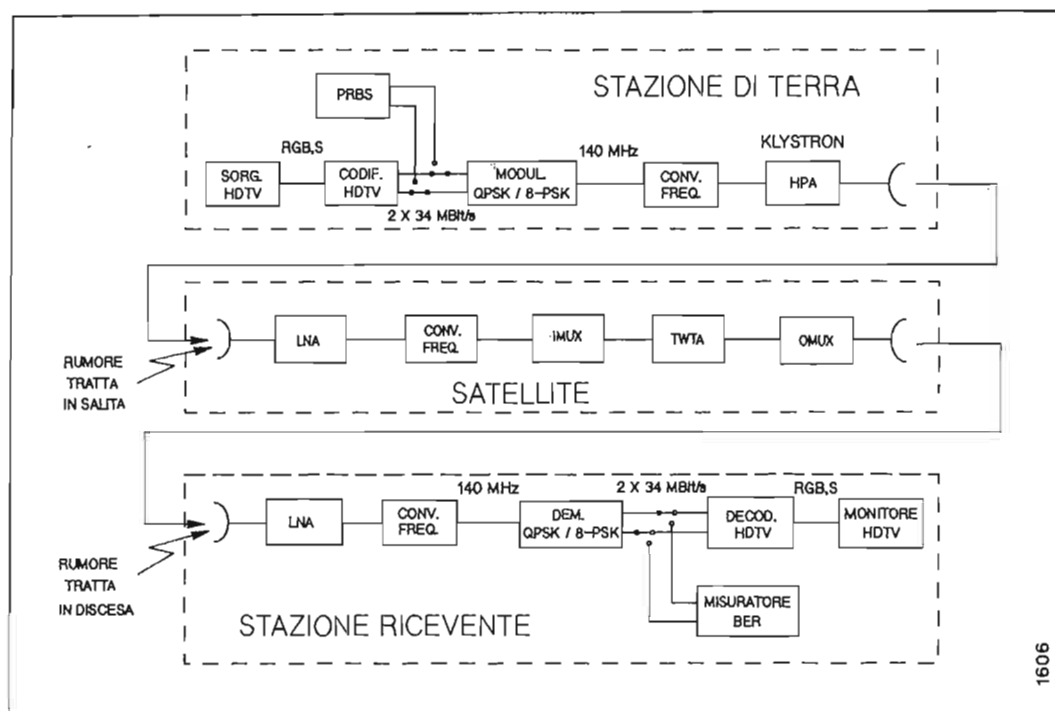



Fig. 1 — Schema a blocchi del canale di trasmissione via satellite.

progressive (TWTA) e un filtro di canale di uscita (OMUX) per ridurre le interferenze sulla tratta in discesa. Normalmente l'amplificatore di bordo viene utilizzato alla saturazione, e quindi in condizioni di forte nonlinearità, per sfruttare appieno la potenza disponibile e migliorare il rendimento energetico. Ciò determina un allargamento dello spettro del segnale numerico irradiato con sensibile risalita dei lobi secondari, analogamente a quanto si può vedere in figura 6 che si riferisce ad un segnale QPSK a 70 Mbit/s all'uscita dell'amplificatore di potenza della stazione trasmittente fissa di Roma.

Il sistema ricevente è composto dall'antenna, da un amplificatore a basso rumore, un convertitore di frequenza e un demodulatore coerente (QPSK o 8-PSK). Il decodificatore HDTV riceve i due flussi numerici a 34,368 Mbit/s in uscita dal demodulatore, li sincronizza recuperando eventuali ritardi di uno rispetto all'altro e ricostruisce i segnali analogici (R-G-B e sincronismi video) per la visualizzazione su monitore o su grande schermo.

Le distorsioni lineari di ampiezza/frequenza e ritardo di gruppo/frequenza della catena trasmissiva generano interferenza intersimbolica sul segnale numerico che si traduce in un peggioramento del tasso di errore del sistema. Come regola molto approssimata le distorsioni che danno maggior contributo sono quelle relative alla banda di Nyquist del segnale modulato. Tale banda è pari a R/2 = 35 MHz ed R/3 = 23,3 MHz, rispettivamente per le modulazioni QPSK e 8-PSK, essendo R = 70 Mbit/s la velocità di trasmissione. Per la modulazione QPSK le variazioni nella risposta ampiezza/frequenza dovrebbero essere contenute entro 0,5 dB e quelle di ritardo di gruppo dovrebbero essere minori di 1/3 del periodo di simbolo Ts (Ts = 28 ns a 70 Mbit/s).

La risposta globale in frequenza (ampiezza e ritardo di gruppo) del canale Olympus (Fig. 2a) è essenzialmente determinata dal filtro di ingresso del satellite, con banda a − 3 dB di circa 42 MHz. Il filtro di uscita è a banda lar-

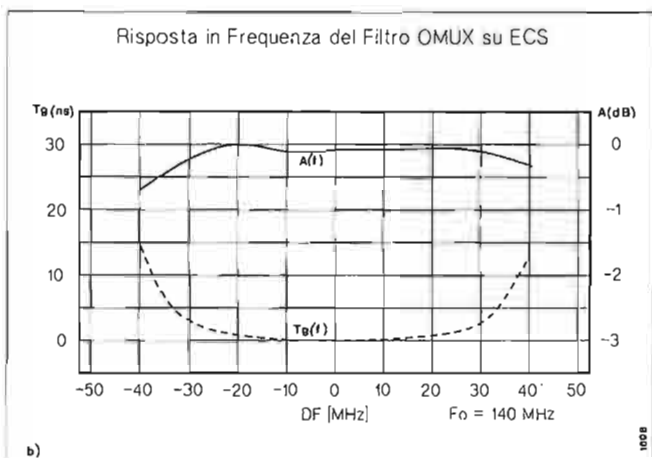
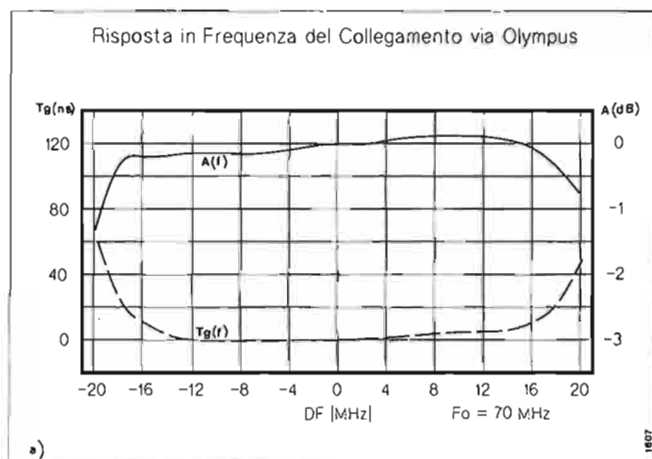



Fig. 2 — Risposta in frequenza del canale di trasmissione *a*) via satellite Olympus e *b*) via satellite EUTELSAT 1.
A(f) = ampiezza / frequenza.
Tg(f) = ritardo di gruppo / frequenza

ga e non protegge quindi i canali adiacenti dalle possibili interferenze introdotte dal segnale numerico il cui spettro, come si è detto, subisce un indesiderato allargamento a causa della non linearità del TWTA. Le caratteristiche del TWTA dell'Olympus (AM/AM e AM/PM) sono riportate in figura 3. Si riscontra un valore massimo del fattore di conversione AM/PM di circa 6°/dB. La stazione trasmittente di terra contribuisce con una distorsione di ampiezza di circa 0,5 dB e di ritardo di gruppo di circa 7 ns a ±20 MHz dalla frequenza centrale del canale.

In figura 2b sono riportate le curve tipiche ampiezza/frequenza e ritardo di gruppo/frequenza del filtro di canale d'uscita dei ripetitori di Eutelsat-I. I filtri d'ingresso hanno una risposta in frequenza molto simile. Considerando gli effetti dei due filtri cumulativi, la banda utilizzabile con basse distorsioni risulta di circa 72 MHz. La banda passante a −3 dB è di circa 80 MHz. Le caratteristiche AM/AM e AM/PM degli amplificatori di potenza (TWTA) di bordo sono simili a quelle dell'Olympus, anche se la potenza è sensibilmente inferiore (20 watt contro 180 watt).

Fig. 3 — Caratteristiche di non linearità dell'amplificatore di potenza (TWTA) a bordo del satellite Olympus (payload DDS, canale 24). AM/AM = potenza in uscita in funzione della potenza in entrata AM/PM = variazione di fase del segnale in uscita in funzione della potenza in entrata

3.3 Prestazioni delle modulazioni PSK via Olympus

3.3.1 *Curve Bit-rate; E_b/N_0; tasso d'errore (BER)*

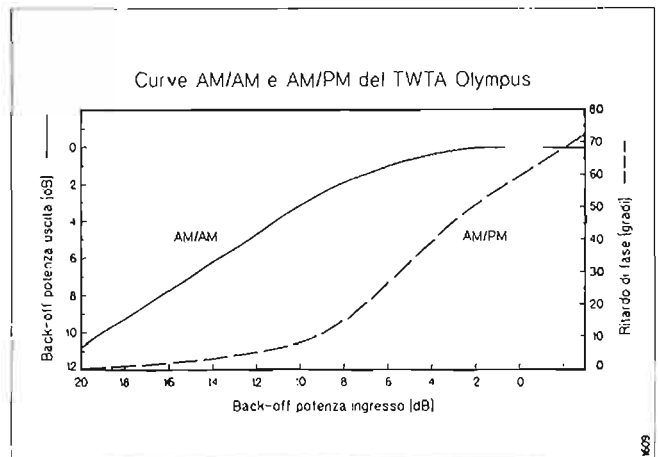
Sul canale Olympus, simulato su calcolatore, sono state valutate le prestazioni delle modulazioni QPSK, Offset-QPSK ed 8-PSK ricavando le curve di tasso di errore (BER) in funzione del rapporto portante-rumore E_b/N_0 (essendo E_b l'energia per bit ed $N_0/2$ la densità spettrale di potenza del rumore). Si ricorda che dal rapporto E_b/N_0 si passa direttamente al rapporto C/N (in 27 MHz) tramite la relazione:

$$C/N\ (dB)_{in\ 27\ MHz} = E_b/N_0\ (dB) + 10\ Log\ R\ (Mbit/s)/27$$

Il rapporto C/N viene normalmente valutato in 27 MHz che è la larghezza di banda normalizzata dalla WARC '77 per i ricevitori analogici TV-MF.

L'Offset-QPSK è una variante del QPSK che presenta minori variazioni di ampiezza del segnale a scapito però di una efficienza spettrale leggermente inferiore.

In figura 4 sono riportati i valori di E_b/N_0 richiesti sul canale Olympus per un BER = 10^{-4} in funzione del bit-rate trasmesso: la modulazione QPSK presenta prestazioni lievemente migliori dell'Offset-QPSK, e può operare senza sensibile degradamento fino a un bit-rate di 60 Mbit/s, mentre a 70 Mbit/s si riscontra un degradamento addizionale di circa 1,2 dB dovuto essenzialmente alle distorsioni lineari del filtro IMUX. Con l'8-PSK, per ottenere lo stesso tasso di errore si richiede un valore di E_b/N_0 maggiore di circa 5 dB, ma è possibile utilizzare bit-rate di 80-90 Mbit/s.

Dai risultati delle simulazioni e dalle prove di laboratorio sui modem disponibili si è deciso di utilizzare la modulazione QPSK, per la sua robustezza alle distorsioni e al rumore e per la maggior affidabilità dei modem disponibili. Il bit-rate scelto è stato di 70 Mbit/s (a livello di modem), il massimo consentito da Olympus con distorsioni ragionevoli, suddiviso in due flussi indipendenti con bit-rate standard di 34,368 Mbit/s forniti dai codec HDTV. Questa scelta consente di ottenere un'ottima qualità delle immagini HDTV co-decodificate, favorisce l'utilizzo di una vasta gamma di dispositivi già esistenti (multiplatori, trasmettitori e ricevitori per fibra ottica) e consente l'accesso alle reti numeriche di terra su due canali a 34 Mbit/s.

In figura 5a sono riportate le curve di BER in funzione di E_b/N_0 ottenute al calcolatore, con prove di laboratorio sul simulatore del satellite e nelle trasmissioni via Olympus a 70 Mbit/s in QPSK attraverso la stazione trasmittente di Roma. Come si vede dalla curva D relativa alle misure via Olympus per ottenere un tasso di errore di 10^{-6} (immagini HDTV non degradate da errori) è richiesto un valore di E_b/N_0 di circa 15 dB, a cui corrisponde un rapporto C/N (in 27 MHz) di circa 19 dB. Durante le prove la stazione trasmittente di terra operava in con-

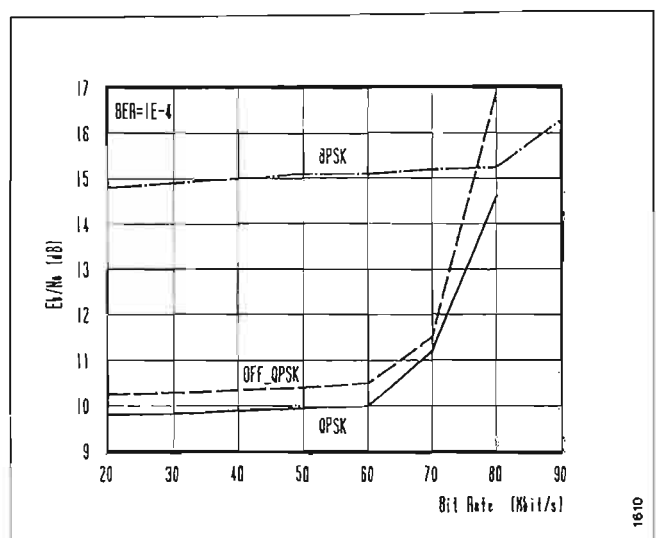



Fig. 4 — Prestazioni delle modulazioni numeriche QPSK, Offset-QPSK ed 8-PSK sul canale di trasmissione via Olympus simulato al calcolatore: Rapporto E_b/N_0 richiesto per un tasso di errore (BER) = 10^{-4} in funzione del bit-rate (Mbit/s).

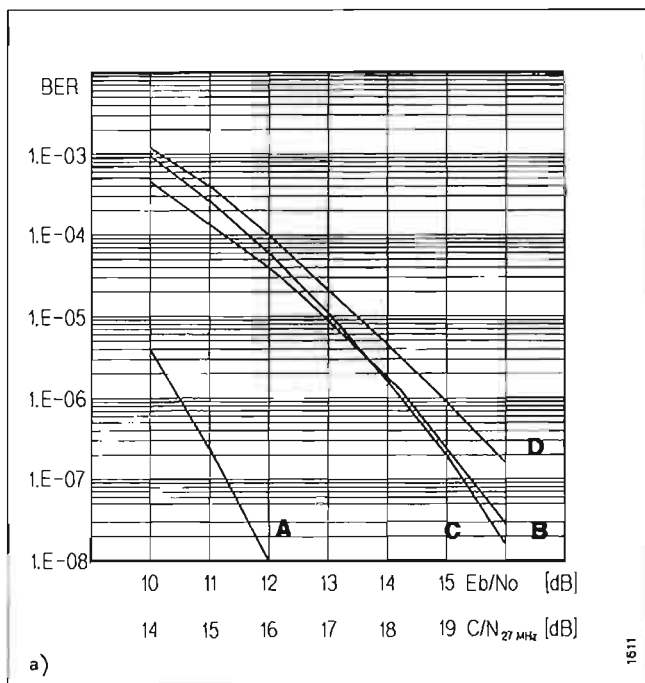



Fig. 5a — A: QPSK ideale; B: simulazioni al calcolatore; C: 70 Mbit/s su simulatore di satellite; D: 70 Mbit/s via Olympus (can. 24)

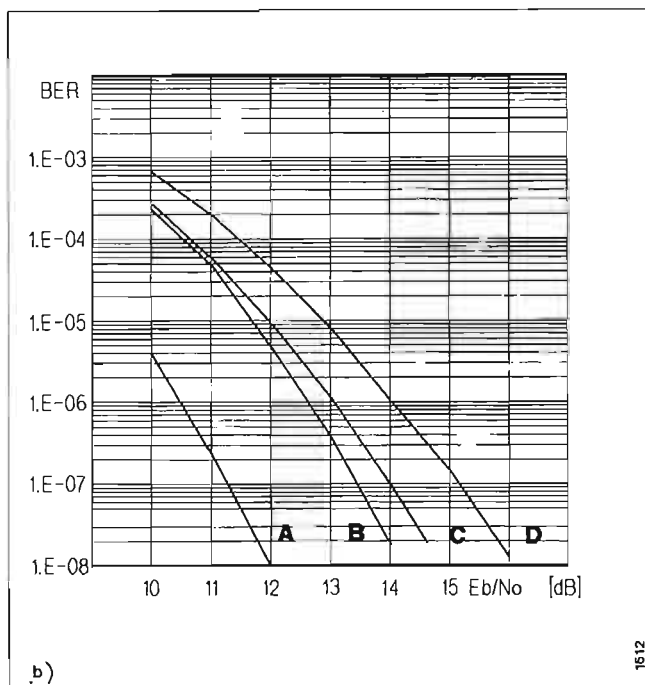



Fig. 5b — A: QPSK ideale; B: 70 Mbit/s su simulatore di satellite; C: 34 Mbit/s su satellite EUTELSAT I; D: 70 Mbit/s su satellite EUTELSAT I

Fig. 5 — Trasmissione a 70 Mbit/s in QPSK *a)* su satellite Olympus e *b)* su satellite EUTELSAT I: tasso di errore (BER) in funzione del rapporto C/N in 27 MHz.

dizioni quasi-lineari al di sotto della potenza massima (con OBO > 3 dB) per non degradare il segnale numerico.

3.3.2 *Riduzione delle interferenze*

La trasmissione di segnali numerici ad elevato bit-rate via Olympus nella banda dei 12 GHz richiede di rispettare precisi vincoli sulle interferenze verso i canali adiacenti stabiliti dalla WARC '77 per il down -link e dall'ORB '88 per l'up-link. In figura 6 sono riportati gli spettri del segnale QPSK a 70 Mbit/s all'uscita della stazione trasmittente di Roma al variare del back-off (OBO) dell'amplificatore a Klystron e sono messe in evidenza le posizioni dei canali adiacenti (AC) a 19,18 MHz (1° AC, interferito dal primo lobo dello spettro QPSK) e a 38,36 MHz (2ª AC, interferito parzialmente dal primo e dal secondo lobo dello spettro). La pianificazione, definita per segnali TV analogici a modulazione di frequenza (MF), prevede i seguenti rapporti di protezione (rapporti fra la potenza del segnale utile e la somma delle potenze dei canali interferenti):

— cocanale: 40 dB (up-link) e 31 dB (down-link)
— canali adiacenti: 21 dB (up-link) e 15 dB (down-link)

Non sono previsti margini di protezione verso i canali a distanza di ± 38,36 MHz (2° AC). Questi canali possono venire invece interferiti da segnali numerici ad elevato bit-rate a causa dell'allargamento dello spettro. Utilizzando i rapporti di protezione sopra indicati, un determinato segnale (analogico o numerico) è considerato compatibile con la pianificazione quando il suo contributo interferenziale nei confronti di un segnale analogico MF è sotto la soglia di visibilità.

Mediante prove di laboratorio (bibl. 2) è stato possibile verificare che le interferenze di un segnale numerico QPSK od 8-PSK su un segnale analogico MF (PAL o MAC) danno origine a un degradamento soggettivo di qualità video simile a quello prodotto da un rumore gaussiano di pari potenza (misurata nel filtro di ricezione da 27 MHz del segnale PAL o MAC interferito). Tale degradamento è molto meno fastidioso di quello prodotto dai segnali video analogici MF di pari potenza che si manifesta in genere sotto forma di battimenti. La soglia di visibilità delle interferenze di un segnale numerico sui segnali TV analogici MF viene raggiunta per valori di C/I (C = potenza del segnale utile, I = potenza del segnale interferente misurata in 27 MHz) dell'ordine di 19 ÷ 21 dB per segnali MAC e di 22 ÷ 24 dB per segnali PAL.

Dati gli elevati rapporti di protezione co-canale previsti dalla pianificazione, le interferenze co-canale da parte di segnali numerici possono quindi essere trascurate.

Per quanto riguarda la *tratta in salita*, con valori di Eirp da terra di 80 dbW (a cui corrisponde un OBO> 5 dB per le stazioni trasmittenti di Italia '90) si sono stimati valori di C/I di circa 30 dB sul primo canale adiacente (a ± 19,18 MHz) e di circa 27 dB sul secondo canale adiacente (a ± 38,36 MHz), molto inferiori quindi alla soglia di visibilità delle interferenze sul MAC e sul PAL. È opportuno rilevare che durante la sperimentazione di Italia '90 i primi canali adiacenti sulla tratta in salita non erano occupati, mentre i secondi adiacenti erano occupati dai programmi in D2-MAC del satellite tedesco TVSAT-2.

Per quanto riguarda la *tratta in discesa*, i valori di C/I stimati sui canali adiacenti sono meno favorevoli rispetto

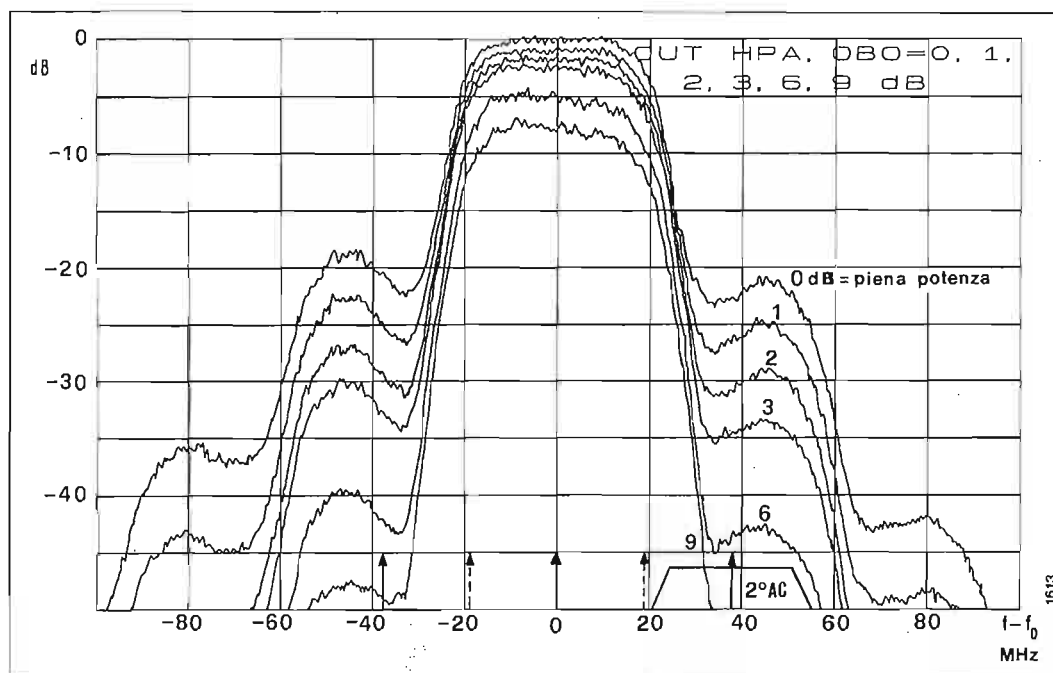



Fig. 6 — Spettro di potenza del segnale QPSK a 70 Mbit/s rilevato all'uscita della stazione trasmittente di Roma per vari valori della potenza irradiata. Con OBO = O dB si ha la potenza massima e con OBO = 1 ± 9 dB si hanno potenze ridotte.

alla tratta in salita, a causa della non-linearità dell'amplificatore di bordo (TWTA) dell'Olympus operante in saturazione, della elevata Eirp del satellite, inferiore di solo 1 dB rispetto ai 64 dBW previsti dalla WARC '77, e dei bassi rapporti di protezione previsti dal piano (15 dB per 1° AC, nessuna protezione per 2° AC). I valori stimati di C/I sono stati di 21 dB sul primo canale adiacente e di 18,5 dB sul secondo canale adiacente, quindi di poco al di sopra della soglia di visibilità per immagini MAC e PAL particolarmente critiche.

Come termine di confronto sono stati valutati con simulazioni al calcolatore i valori di C/I relativi al segnale C-MAC normalizzato dal CCIR (modulazione 2-4 PSK a 20,25 Mbit/s), con multiplex audio/dati a pieno canale, interferente su un segnale TV analogico MF nelle seguenti condizioni:

— TWTA del satellite operante in saturazione
— filtro d'uscita (OMUX) a banda larga
— Eirp massima prevista dalla WARC '77 (64 dB watt)

Si sono ottenuti valori di C/I di circa 21,5 dB sul 1° AC e di 26 dB sul 2° AC. Dunque, mentre sul primo canale adiacente si verifica circa lo stesso livello interferenziale di un segnale QPSK a 70 Mbit/s, sul secondo la situazione è molto migliore. Per ottenere con modulazione QPSK valori di C/I di 26-27 dB sul secondo canale adiacente è necessario ridurre il bit-rate a circa 50 ÷ 60 Mbit/s ed utilizzare filtri OMUX del tipo impiegato su TV-SAT (50 MHz di banda). Per soddisfare le maschere sullo spettro irradiato previste per il C-MAC (bibl. 3) qualora si trasmetta con segnali numerici in QPSK è necessario impiegare un bit-rate massimo di circa 45 Mbit/s.

Durante Italia '90 i canali adiacenti non erano occupati e quindi il problema interferenziale sulla tratta in discesa non si poneva. In caso contrario per trasmettere a 70 Mbit/s sarebbe stato necessario scegliere una modulazione a banda più stretta e ridurre leggermente la potenza irradiata dal satellite per far lavorare il TWTA in condizioni più lineari: ad esempio i rapporti C/I ottenibili su Olympus con modulazione 8-PSK a 70 Mbit/s (assumendo una riduzione di potenza del TWTA di circa 1 dB, senza filtro di canale in uscita) risultano di circa 24 dB (1° AC) e 27 dB (2° AC), migliori quindi di quelli ottenibili con il C-MAC. In queste condizioni di lavoro lo spettro dell'8-PSK a 70 Mbit/s all'uscita del satellite, contrariamente al QPSK alla stessa velocità, soddisfa pienamente la maschera definita per il C-MAC. Tuttavia, data la maggiore sensibilità al rumore, per operare con un tasso di errore di 10^{-6} sul canale l'8-PSK a 70 Mbit/s si richiede un rapporto C/N di 23 dB, superiore quindi di 4 dB al valore di 19 dB richiesto dalla modulazione QPSK allo stesso bit-rate. Per soddisfare tale esigenza sarebbe necessario aumentare di circa il 60% il diametro delle antenne riceventi.

3.4 Prestazioni delle modulazioni PSK via Eutelsat I

I satelliti Eutelsat e Intelsat sono correntemente utilizzati per telefonia/dati con modulazioni numeriche QPSK a 120 Mbit/s con accesso a multiplazione di tempo (TDMA). Data la limitata potenza irradiata dai satelliti, le stazioni rice-trasmittenti di terra richiedono antenne di grosso diametro (almeno 8 metri).

Le prestazioni del QPSK e dell'8-PSK, in termini di BER in funzione di E_b/N_o, sono simili a quelle ottenibili su Olympus, ma la maggiore larghezza di banda del ripetitore (80 MHz contro 42 MHz) permette di aumentare il bit-rate di un fattore di circa 1,9.

In figura 5b sono riportate le curve di BER in funzione di E_b/N_o ottenute sul simulatore di satellite e con trasmissioni via EUTELSAT-I (F4) dalla stazione di up-link di Telespazio, al Fucino, nelle seguenti condizioni:

— modulazione QPSK a 34 a 70 Mbit/s
— trasmissione e ricezione con antenna da 8 metri
— Eirp da terra di 77,6 dBW (HPA operante in quasi-linearità, OBO = 10,4 dB)
— TWTA del satellite in saturazione (OBO = O dB)

Le curve sono state ottenute sommando rumore all'entrata del demodulatore QPSK. Mentre le prestazioni a 34 Mbit/s sono simili a quelle ottenute sul simulatore del satellite, quelle a 70 Mbit/s presentano un degradamento aggiuntivo di circa 1,5 dB (a BER = 10^{-6}), probabilmente imputabile alle distorsioni introdotte dal ricevitore (progettato per segnali TV/MF), e in parte recuperabile con un ricevitore ottimizzato per i segnali numerici. In ogni caso il valore di E_b/N_0 richiesto per BER = 10^{-6} è compreso tra i 13 e i 14 dB (contro i 15 dB richiesti via Olympus) a cui corisponde un valore di C/N in 27 MHz di 17 ÷ 18 dB per trasmissioni HDTV a 70 Mbit/s in QPSK.

Per quanto riguarda le interferenze sulla tratta in salita, le misure effettuate indicano che per rispettare le maschere fissate da Eutelsat si deve far lavorare l'amplificatore a Klystron della stazione di terra con un OBO di almeno 2 ÷ 3 db.

3.5 Bilanci di collegamento e scelta del satellite

I bilanci di collegamento via Olympus e via Eutelsat (EUTELSAT I/F4), per il dimensionamento delle stazioni trasmittenti e riceventi, sono stati analizzati in dettaglio in (Bibl. 4 e 5). L'ipotesi di base era quella di utilizzare stazioni trasmittenti trasportabili per l'accesso diretto al satellite dagli stadi di Italia '90, in modo da evitare i costi e la complicazione operativa relativa al trasporto e alla distribuzione del segnale HDTV su reti numeriche di terra. Di seguito sono riportati i risultati principali di queste valutazioni che, nel caso dell'Olympus, tengono conto dei dati sul satellite forniti dall'ESA (G/T, Eirp, ecc.) dopo le prove in orbita.

3.5.1 *Satellite Olympus, canale 24*

Per la tratta in salita a 18 GHz, nell'ipotesi di utilizzare una stazione trasportabile da 85 dBW (Eirp massima), operante con una Eirp di circa 80 ÷ 82 dBW (OBO = 5 – 3 dB) per ridurre le distorsioni non-lineari sul segnale numerico e le interferenze sui canali adiacenti, il C/N in 27 MHz stimato è risultato di circa 25,3 ÷ 27,3 dB (statistica dell'attenuazione da pioggia al 99% del mese peggiore dell'anno, area di copertura a –2 dB rispetto al centro del fascio Europeo, puntato tra l'Italia Settentrionale e la Svizzera).

Grazie ad un sofisticato sistema di controllo di guadagno il satellite opera con il TWTA alla saturazione per tutti i valori di Eirp di interesse sulla tratta in salita. Il C/N (in 27 MHz) sulla tratta in discesa stimato disponibile al 99% del mese peggiore è risultato di 26,6 dB con G/T della stazione ricevente di 20,7 dB/°K (diametro di antenna di 1,8 m, G = 44,7 dB, cifra di rumore dell'amplificatore a basso rumore NF = 1,7 dB).

Con i valori di C/N dedotti per le tratte in salita e in discesa il C/N totale disponibile è risultato di 22,9 ÷ 23,9 dB, a seconda della potenza irradiata dalla stazione trasmittente di terra, con un margine di 3,9 ÷ 4,9 dB rispetto ai 19 dB richiesti per il corretto funzionamento del sistema.

3.5.2. *Satellite EUTELSAT I/F4*

In condizioni operative a cielo chiaro si richiede una Eirp da terra di 77,6 dBW per ottenere la massima potenza all'uscita del satellite (TWTA alla saturazione). Per poter effettuare le trasmissioni da una stazione trasportabile, con antenna di diametro 3 ÷ 4 metri, è stato ipotizzato l'utilizzo del satellite con «gain setting» = 20, operante cioè nella condizione di massimo guadagno. Ciò permettere di far operare il TWTA di bordo a piena potenza con i valori di Eirp disponibili dalle stazioni di terra. Ottimizzando i parametri delle stazioni rice-trasmittenti (risposta in frequenza, rumore di fase dei convertitori, ecc.) si può stimare un valore di C/N (in 27 MHz) di circa 17 ÷ 18 dB per un tasso di errore sul canale di 10^{-6} (immagini HDTV non degradate da errori di trasmissione).

I bilanci di collegamento sono stati effettuati nelle seguenti ipotesi: tratta in salita in Eurobeam, trasmissione da Torino, ricezione nello spot Ovest (con centro sulla Manica), stazione trasportabile da 77 dBW di Eirp alla saturazione con antenna da 4 m equipaggiata di step-track (stazione non disponibile in RAI), cifra di rumore del ricevitore NF = 1,8 dB. Trasmettendo con 75 dBW (2 dB di output back-off per rispettare le maschere Eutelsat sulle interferenze) si ottiene un C/N (in 27 MHz) di circa 17 dB (ricezione a Roma) a cielo chiaro impiegando antenne riceventi da 4 m. In tali condizioni si verrebbe quindi ad operare al limite delle prestazioni del sistema numerico, senza margini di ricezione. Le condizioni di ricezione nell'Italia meridionale risultano sensibilmente più critiche e richiedono antenne riceventi di almeno 5 ÷ 6 m equipaggiate con puntamento automatico. Solo nell'Italia settentrionale sarebbe possibile operare con antenne riceventi di circa 3 m, seppure con scarsi margini di ricezione in cattive condizioni atmosferiche.

L'ipotesi di raggiungere le sale visione con due salti via satellite, il primo tra la stazione trasportabile e la stazione fissa di ricezione di Telespazio al Fucino (equipaggiata con una antenna da 8 metri), il secondo tra il Fucino e le sale visione, avrebbe permesso di operare con «gain setting» del satellite in posizione normale, ma con antenne riceventi di diametro non inferiore a 3 m al nord e con diametri assai maggiori al sud. Anche questa ipotesi è stata scartata, inoltre, per la difficoltà di disporre di due ripetitori dedicati su Eutelsat I. Un'altra seria difficoltà era dovuta alla mancanza di una stazione trasportabile adatta per segnali numerici a 70 Mbit/s QPSK ed in grado di operare con Eirp di almeno 77 dBW.

4. **Risultati della sperimentazione tecnica ad Italia '90**

Sulla base dei risultati della fase preparatoria, che evidenziavano i numerosi problemi relativi all'utilizzazione del satellite Eutelsat I, l'ipotesi di ricorrere all'Olympus diventava prioritaria, e si affermava definitivamente dopo il successo delle prove di trasmissione a 70 Mbit/s dal-

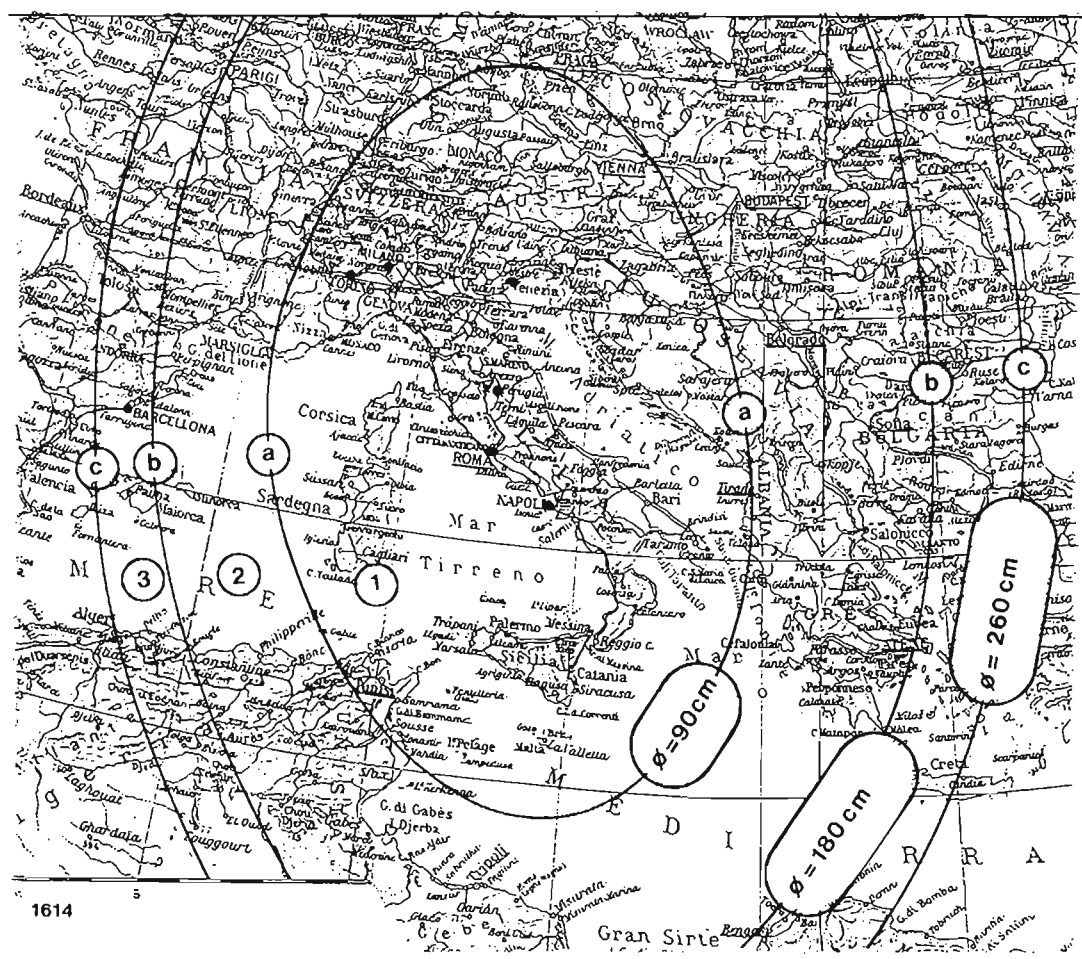



Fig. 7 — Nella figura sono indicati i punti riceventi e sono individuate tre zone: 1, 2 e 3, delimitate da tre curve: a, b e c, sulle quali sono indicati i diametri delle antenne riceventi necessari, all'interno delle zone, per ottenere un rapporto C/N di 19 dB (in 27 MHz) per il 99% del tempo nel mese peggiore.

la stazione fissa di Roma (bibl. 6). L'uso dell'Olympus era inoltre vantaggioso per vari motivi:

— la completa disponibilità del canale 24, normalmente utilizzato per le trasmissioni sperimentali di RAI-SAT;
— la grande area di copertura consentita dall'elevata potenza irradiata (Eirp = 63 dBW sull'asse dell'antenna trasmittente);
— la disponibilità della stazione fissa di feeder-link, presso il Centro di produzione RAI di Roma e di una stazione trasportabile in costruzione presso la Selenia Spazio.

La flessibilità operativa dei codec Telettra ha consentito di operare con due catene di produzione HDTV separate: una secondo il formato di scansione europeo a 1250 linee/50 Hz gestita da EU-95 con l'appoggio diretto della RAI ed installata presso lo stadio Olimpico di Roma. I segnali di banda base video ed audio (un canale stereo ed un canale con effetti sourround) forniti dai mezzi di ripresa venivano codificati sotto forma numerica ed inviati tramite collegamento in fibra, su due flussi a 34,368 Mbit/s, al Centro di produzione RAI per la trasmissione verso l'Olympus a 70 Mbit/s in QPSK sul canale 24. Gli stessi segnali di banda base venivano anche inviati ad un codificatore HD-MAC installato da EU-95 presso lo stadio Olimpico. Il segnale HD-MAC veniva quindi inviato su collegamento in fibra ottica al Centro di produzione RAI per essere trasmesso con modulazione di frequenza sul canale 20 dell'Olympus (il canale a copertura europea normalmente utilizzato dalla BBC) verso alcuni punti di ricezione dislocati in Europa. Questa operazione era gestita diretamente da EU-95. La stazione trasmittente fissa della RAI, a Roma, trasmetteva quindi contemporaneamente in diretta via Olympus le cinque partite giocate allo stadio Olimpico in HDTV numerica a 70 Mbit/s sul canale 24 ed in HD-MAC sul canale 20.

La seconda catena di produzione operava nel formato di scansione a 1125 linee/60 Hz con un automezzo di ripresa dell'NHK direttamente interconnesso alla stazione trasmittente trasportabile del *Centro Ricerche* RAI sulla quale erano installati il codec HDTV ed il modem a 70 Mbit/s QPSK. L'agilità operativa di questa seconda catena di produzione consentiva alla RAI di riprendere in diretta 10 partite giocate negli stati di Torino, Milano, Firenze e Napoli.

Otto punti riceventi, che alimentavano le sale di visione HDTV con proiezione su grande schermo, venivano installati dal *Centro Ricerche* RAI a Torino (2), Milano, Venezia, Perugia, Roma (2) e Napoli. La configurazione della rete di trasmissione e distribuzione del segnale via Olympus è descritta in dettaglio in bibliografia 7.

La figura 7 mostra l'area di copertura dell'Olympus sul canale 24. Sono indicate le posizioni delle installazioni riceventi realizzate dal *Centro Ricerche* in Italia e l'installazione realizzata dalla Retevision a Barcellona. Qui il segnale ricevuto dal satellite veniva inviato a Madrid, su due flussi a 34 Mbit/s, tramite un collegamento in fi-

bra ottica di circa 620 Km. A Barcellona ed a Madrid erano allestite due sale visione con proiezione su grande schermo.

È indicato sulle varie curve di figura 7 il diametro dell'antenna ricevente necessario per ottenere per il 99% del mese peggiore un rapporto C/N (in 27 MHz) di almeno 19 dB, assumendo una figura di rumore del ricevitore di 2 dB. Il valore di 19 dB del rapporto C/N corrisponde al *minimo* richiesto per il corretto funzionamento del sistema (vedi par. 3.3), ossia con immagini HDTV non degradate da errori di trasmissione (BER $< 10^{-6}$ sul canale). Per garantire tale valore di C/N nei paesi situati sulla curva a -3 dB di figura 7 si richiedono antenne da 90 cm mentre a Barcellona occorre una antenna di almeno 2 m.

Data l'importanza della sperimentazione, in cui veniva coinvolto un vasto pubblico, si decideva di aumentare sensibilmente il diametro delle antenne riceventi onde garantire adeguati margini di ricezione anche in condizioni di propagazione molto sfavorevoli (ad es. in presenza di pioggia sia sulla tratta in salita a 18 GHz che su quella in discesa a 12 GHz). Nei punti riceventi in Italia si impiegavano antenne da 1,8 m mentre a Barcellona veniva installata una antenna da 2,8 m.

Per quanto riguarda le prestazioni dell'intero collegamento numerico che, oltre all'impiego del satellite Olympus, includeva tratte su fibra ottica si sono ottenuti i seguenti risultati:

— i valori del rapporto C/N misurati a cielo chiaro nelle varie località sul territorio nazionale risultavano compresi tra 29 dB al nord (Torino e Milano) e 32 dB a Roma. Questi valori sono risultati in buon accordo con le stime ottenute dai bilanci di collegamento;
— gli elevati margini di ricezione del segnale (10 dB al nord e 13 dB a Roma) hanno consentito di operare con una disponibilità del servizio superiore al 99,9% durante l'intero periodo delle dimostrazioni, per un totale di circa 50 ore di trasmissione;
— il segnale HDTV è stato ricevuto con successo a Torino, a scopi sperimentali, anche con un'antenna di soli 70 cm, con un margine di C/N di circa 3 dB a cielo chiaro;
— i sistemi riceventi appositamente sviluppati dal Centro Ricerche si sono dimostrati affidabili e perfettamente adeguati alla ricezione del segnale numerico;
— la stazione trasmittente installata presso il Centro di produzione RAI di Roma ha operato con stabilità ed affidabilità a lungo termine senza richiedere particolari interventi di allineamento e controllo; qualche problema, peraltro sempre superato dal personale addetto, si è ayuto con la stazione trasmittente trasportabile che si è spesso trovata ad operare in difficili condizioni ambientali. Si deve rilevare però che tale stazione è entrata praticamente in servizio al momento stesso della consegna da parte della Selenia Spazio, senza la possibilità di completare le operazioni di allineamento e collaudo.
— i collegamenti in fibra ottica si sono dimostrati totalmente «trasparenti» al trasporto del segnale numerico;

È importante osservare che le prestazioni dell'intero sistema di trasmissione a 70 Mbit/s in QPSK possono essere ulteriormente migliorate a breve termine operando sia sul modem che sul codec HDTV. Sul modem è possibile infatti impiegare uno scrambler-descrambler di tipo «sincronizzato» sulla trama in sostituzione a quello di tipo «autosincronizzante» disponibile ad Italia '90. Ciò consentirebbe di operare con un tasso di errore sul canale di 10^{-5}, invece di 10^{-6}. Un guadagno di circa 1,5 dB sul C/N sarà possibile con la seconda generazione di codec HDTV che disporranno di un codice correttore di errore di tipo RS(255,239), sensibilmente superiore all'attuale codice BCH(255,239). Con questi possibili miglioramenti sul modem e sul codec il sistema potrà operare via Olympus con un rapporto C/N minimo di circa 16 dB contro i 19 dB di Italia '90. È importante osservare inoltre che l'impiego di tecniche di «compensazione del movimento» nei codec della seconda generazione consentirà di migliorare ulteriormente la qualità delle immagini a parità di bit-rate (70 Mbit/s), oppure di ridurre ulteriormente il bit-rate (presumibilmente verso i 50 ÷ 60 Mbit/s) mantenendo la stessa qualità dimostrata ad Italia '90.

In un'ottica a più lungo termine sarà possibile disporre di tecniche di modulazione e codifica di canale più sofisticate basate sulla decodifica di Viterbi, in grado di migliorare sensibilmente le prestazioni dell'intero collegamento. A titolo di esempio la modulazione Trellis-coded 8-PSK con rate 5/6 (bibl. 8) consentirà di eliminare le distorsioni lineari introdotte dal filtro IMUX dell'Olympus e ridurre le interferenze sui canali adiacenti (grazie ad una occupazione spettrale ridotta del 20% rispetto al QPSK), oltre a consentire un certo guadagno di codifica (1 ÷ 2 dB). La stessa modulazione con rate 2/3, o la modulazione QPSK associata a codifica convoluzionale con rate 3/4 (bibl. 9), entrambe concatenate ad un codice di Reed-Solomon (255,239) associato al codec HDTV, consentirebbero di operare anche su Eutelsat I con antenne riceventi di diametro contenuto. I valori del rapporto C/N in 27 MHz richiesti da questi sistemi, includendo 2 dB di margine di implementazione, sono di circa 13 dB (per il Trellis coded 8-PSK rate 2/3) e 11,5 dB (per il QPSK con convoluzionale 3/4), contro i 16 ÷ 17 dB richiesti con la semplice modulazione QPSK ed il codice BCH(255,239).

5. Conclusioni

Gli esperimenti effettuati dalla RAI nell'occasione della Coppa del Mondo di calcio (Italia '90) hanno dimostrato, in anteprima mondiale, la fattibilità tecnica ed operativa di un sistema di trasmissione numerica a 70 Mbit/s per il trasporto punto-multipunto di segnali HDTV di alta qualità tramite una articolata rete di distribuzione includente il satellite Olympus e collegamenti numerici in fibra ottica, anche su lunga distanza.

Le principali conclusioni che si traggono dalla sperimentazione si possono così sintetizzare:

— la qualità delle immagini ricevute da satellite, visualizzate su monitori HDTV di tipo professionale, è risultata potenzialmente uguale a quella disponibile alla sorgente. Il sistema di co-decodifica a 70 Mbit/s sviluppato nell'ambito del progetto europeo EU-256 si è dimostrato quindi «trasparente» per la distribuzione di segnali HDTV di elevata qualità;
— le elevate prestazioni dell'intera catena di trasmissio-

ne numerica, in termini di immunità agli errori di trasmissione con elevati margini di corretta ricezione anche in cattive condizioni di propagazione, hanno confermato i vantaggi della soluzione numerica nelle applicazioni punto-multipunto. A questo si aggiunga la possibilità dei codec numerici di operare in un contesto di produzione HDTV misto, a 50 Hz ed a 60 Hz. I risultati conseguiti ad Italia '90 consentiranno alla RAI di contribuire alla futura attività del Project Group 04 di EU-95 che prevede lo sviluppo di sistemi numerici avanzati per il trasporto punto-punto via satellite di segnali HDTV con qualità contributo.
Sarà inoltre possibile contribuire agli studi dell'UER e del CCIR sui sistemi di HDTV numerici a «larga banda» per diffusione diretta da satellite in vista della WARC '92, nonché all'attività del CNR nel quadro del Piano Finalizzato delle Telecomunicazioni.

(3930)

RINGRAZIAMENTI

Gli Autori desiderano ringraziare l'Ing. G. P. Pacini e l'Ing. G. Moro per la collaborazione nella valutazione dei link-budget e nel dimensionamento delle stazioni di terra per Olympus ed Eutelsat I.

BIBLIOGRAFIA

1. - M. Cominetti, A. Morello: *Trasmissione da satellite di segnali televisivi numerici: prestazioni e prospettive*, 35° Congresso per L'Elettronica, Roma 1988.

2. - M. Cominetti, A. Morello, M. Visintin: *Digital transmission of television signals by satellite*, ITU COM '89, Geneva, October 1989.

3. - EBU: *Specification of the Systems of the MAC/packet family*, Doc. Tech. 3258, 2ª Edizione, EBU Bruxelles, Luglio 1990.

4. - G. P. Pacini: *Ipotesi di trasmissione HDTV a 70 Mbit/s tramite satellite EUTELSAT dagli stadi. Link-budget*, «RAI-Centro Ricerche», Relazione Tecnica n. 88/13/I.

5. - G. P. Pacini: *Stazione mobile per trasmissione a Olympus. Pacco 18-12 GHz*, «RAI-Centro Ricerche», Relazione Tecnica n. 88/14/I.

6. - G. Garazzino, A. Morello, B. Sacco: *Prove di trasmissione di segnali QPSK e 70 Mbit/s via Olympus dalla stazione del CPTV RAI di Roma*, «RAI-Centro Ricerche», Relazione Tecnica n. 90/01/I.

7. - M. Ardito, G. F. Barbieri, M. Cominetti: *ITALIA '90: Prima mondiale di collegamento HDTV numerico via satellite*, «Elettronica e Telecomunicazioni» in questo numero.

8. - G. Montorsi, A. Morello: *Performance of trellis coded 8-PSK with rate 2/3 on satellite channels*, «RAI-Centro Ricerche», Relazione Tecnica no. 90/2/E.

9. - M. Visintin: *Performance of QPSK with rate 1/2 and 3/4 convolutional codes on satellite channels*, «RAI-Centro Ricerche», Relazione Tecnica no. 90/1/E.

ITALIA '90: CODIFICA DEL SEGNALE TELEVISIVO NUMERICO

M. BARBERO, S. CUCCHI, R. DEL PERO, G. DIMINO,
M. OCCHIENA, M. MURATORI, M. STROPPIANA*

SOMMARIO — Il co-decodificatore utilizzato nel corso dell'esperimento di trasmissione punto-multipunto via satellite di segnali HDTV in occasione dei campionati mondiali di calcio Italia '90 è stato sviluppato nell'ambito del progetto europeo Eureka EU 256. Su «Elettronica e Telecomunicazioni» sono già comparsi numerosi articoli relativi alle attività di ricerca volte all'ottimizzazione di tale sistema di riduzione della ridondanza, sistema basato sull'uso della Trasformata Coseno Discreta (DCT). Il presente articolo è corredato di un'ampia bibliografia e riassume brevemente le tappe fondamentali di tali studi e dell'attività di normalizzazione internazionale per quanto riguarda i sistemi di compressione del segnale TV e HDTV. Descrive inoltre la struttura del codec HDTV impiegato durante Italia '90 e fornisce indicazioni sulle attività di ricerca e di sviluppo previste per i prossimi mesi.

SUMMARY — **Italia '90: Digital TV signal coding** *- The co-decoder used in the experiments of point-to-multipoint satellite transmission of HDTV signals, on the occasion of the FIFA WorldCup Italia '90, has been developed in the Framework of the European project Eureka EU 256. Various articles about the research activities to optimise this bit-rate reduction system, based on the Discrete Cosine Transform (DCT), have already been issued on «Elettronica e Telecomunicazioni». This article is provided with an extended list of references and briefly summarises the main steps of such studies and the international activities to standardise the compression system to be applied to TV and HDTV signals. Morevore it describes the actual structure of the HDTV codec used for Italia '90 and provides some information on the research and development activities foreseen for the next few months.*

1. Premessa

Un elemento fondamentale per la riuscita dell'esperimento di collegamento punto-multipunto realizzato nel corso di Italia '90 è stato ovviamente il co-decodificatore. La sua funzione è quella di ridurre la ridondanza presente nel segnale HDTV al fine di consentirne la trasmissione utilizzando un canale di capacità limitata, quale quello disponibile tramite il satellite Olympus.

Sul sistema di riduzione della ridondanza sviluppato nell'ambito del progetto europeo Eureka EU 256[1] è già comparsa una serie di articoli su questa rivista.

Per tale ragione in qualità di autori degli articoli già pubblicati o di prossima pubblicazione ci limitiamo, in occasione di questo numero speciale di Elettronica e Telecomunicazioni, a riassumere brevemente le caratteristiche principali del sistema EU 256 e a fornire alcune indicazioni sugli ulteriori sviluppi previsti per i prossimi mesi.

(*) Ing. Marzio Barbero, ing. Roberto Del Pero, ing. Giorgio Dimino, p.i. Massimo Occhiena, ing. Mario Muratori, dott. Mario Stroppiana del Centro Ricerche RAI - Torino e ing. Silvio Cucchi della Telettra S.p.A. - Vimercate (MI).
Dattiloscritto pervenuto alla Redazione il 29 ottobre 1990.

[1] EU 256 (Bit-Rate Reduction System for HDTV Digital Transmission) è il progetto europeo Eureka tra la RAI e la Telettra S.p.A. per l'Italia e la Retevision, la Telettra Espana S.A. e Universidad Politecnica de Madrid per la Spagna.

2. Alcuni cenni storici

La presentazione dell'attività di normalizzazione a livello internazionale e della ricerca volta a determinare ed ottimizzare la struttura degli algoritmi di codifica, la descrizione del sistema di riduzione della ridondanza e la divulgazione dei risultati via via ottenuti sono state oggetto di numerosi articoli e contributi a convegni (bibl. 1, 2, e «Elettronica e Telecommunicazioni», n. 1, 1989 e n. 1, 1990).

Il sistema di riduzione del flusso di informazione binaria (BRR, Bit-Rate Reduction) studiato nell'ambito di EU 256 è basato sull'uso della Trasformata Coseno Discreta (DCT, Discrete Cosine Transform) che consente di sostituire agli elementi di immagine (pel) i coefficienti DCT che presentano una minore mutua correlazione. Un'analisi dei metodi di riduzione della ridondanza e la formulazione matematica della DCT sono riportati in precedenti articoli. La DCT ha assunto un ruolo fondamentale in tutti i sistemi di BRR e chi è interessato ad una trattazione ampia e aggiornata dell'argomento può fare riferimento ad un libro di recentissima pubblicazione (bibl. 3).

La figura 1, precedentemente pubblicata, presenta una descrizione schematica del sistema di codifica ibrido utilizzato, in grado di sfruttare sia la ridondanza temporale, mediante codifica predittiva, sia quella spaziale, grazie alla trasformata DCT.

Per ridurre la ridondanza statistica associata ai coefficienti DCT relativi alle immagini video (bibl. 4) si fa uso dei codici a lunghezza variabile (VLC, Variable Lenght Code); nei succitati numeri di «Elettronica e Telecomu-

Fig. 1 — Schema a blocchi di un codificatore basato sull'uso della Trasformata Coseno Discreta ibrida. I campioni in ingresso, organizzati in gruppi di 8 · 8, vengono trasformati (DCT), moltiplicati per un fattore di scalamento (SCALER) in base al riempimento del buffer, quantizzati (Q), codificati con parole a lunghezza variabile (VLC) e memorizzati nel BUFFER. Ad essi viene successivamente aggiunta la ridondanza necessaria per la protezione dagli errori introdotti dal canale (FEC). I dati di ingresso al circuito che calcola la DCT possono essere campioni video (modo intra-field) oppure differenze fra il valore di ciascun campione e un valore di predizione ottenuto a partire dai campioni appartenenti al semiquadro precedente (modo inter-field) o da quelli appartenenti al semiquadro ancora prima (modo inter-frame). In quest'ultimo caso i valori di predizione possono essere ottenuti applicando tecniche di stima e compensazione del movimento. I dati di predizione devono essere uguali, a meno di differenze dovute ad errori di trasmissione, a quelli disponibili dal lato di decodifica; per tale ragione i campioni di riferimento memorizzati nella memoria di quadro sono ottenuti mediante le operazioni inverse di quantizzazione (Q^{-1}), scalamento ($SCALER^{-1}$) e trasformata (DCT^{-1}) a partire dai dati in trasmissione.

nicazioni» si descrivono i diversi VLC proposti in sede internazionale e si analizza in particolare il B-code, cioè il VLC utilizzato nel codec EU 256.

La struttura di trama del segnale video codificato, insieme alla scelta del codice VLC, ha una parte rilevante nelle prestazioni del sistema in presenza di errori introdotti dal canale (bibl. 5).

La scelta di un'organizzazione a pacchetti dei dati video è stata inoltre operata al fine di semplificare la realizzazione del codec con le tecnologie attualmente disponibili e consentire l'uso di un'architettura modulare per il codec HDTV (bibl. 6).

Il sistema sviluppato in EU 256 è in grado di operare su segnali video di diverso formato e di adattare il flusso di dati generato alle capacità del canale disponibile.

Una versione del codec è utilizzata per la codifica dei segnali video a componenti e a definizione convenzionali, quelli cioè conformi alla Raccomandazione CCIR 601 e normalmente indicati con il termine «segnali 4:2:2». I parametri di tale sistema di codifica sono stati oggetto di discussione nel gruppo di lavoro interinale IWP CMTT/2, a cui hanno partecipato attivamente alcuni dei ricercatori coinvolti nel progetto EU 256, in rappresentanza di Italia e Spagna. Recentemente è stata emessa una nuova Raccomandazione, la 723, che specifica i principali parametri di codifica per i segnali televisivi a 525 e 625 righe. I risultati ottenuti nel corso del lavoro di ottimizzazione dell'algoritmo e la dimostrazione della loro realizzabilità pratica hanno fatto sì che il metodo di BRR specificato nella Raccomandazione sia sostanzialmente in accordo con le scelte operate per il codec EU 256.

La raccomandazione 723 attualmente non specifica il VLC, la struttura di trama video e i metodi di adattamento alla rete, cioè ai diversi protocolli di trasmissione. Tali elementi sono ancora in discussione: è probabile che prossimamente venga prodotto uno standard europeo da parte dell'ETSI (European Telecommunications Standards Institute) per la trasmissione di segnali 4:2:2 per applicazioni ad uso contributo a 34 Mbit/s. Tale norma sarà basata sui parametri indicati nella 723, ma completa anche per quanto riguarda gli elementi in essa ancora mancanti. L'UER (Union Europeenne de Radiodiffusion) si è impegnata ad utilizzare tale norma, appena disponibile, per la rete eurovisione. Se ciò avverrà, è probabile che anche la 723 venga completata sulla base della scelta operata in ambito europeo e che si pervenga quindi ad un sistema unico mondiale per la trasmissione a 34 e 45 Mbit/s del segnale 4:2:2.

3. La situazione attuale

I codec utilizzati nel corso delle sperimentazioni di Italia '90 sono basati sulla struttura descritta nelle sue linee essenziali in bibliografia 6. Allo scopo di riassumere brevemente le caratteristiche del codificatore HDTV, si riporta in figura 2 lo schema a blocchi relativo alla sua architettura, mentre in figura 3 è rappresentato uno degli apparati di codifica utilizzato nel corso di Italia '90.

La struttura modulare del codificatore e, in modo del tutto duale, del decodificatore consente di dividere il flusso binario corrispondente ad un segnale HDTV in quattro flussi, ciascuno corrispondente ad un segmento verticale dell'immagine da elaborare.

Si consideri, ad esempio, il segnale corrispondente al

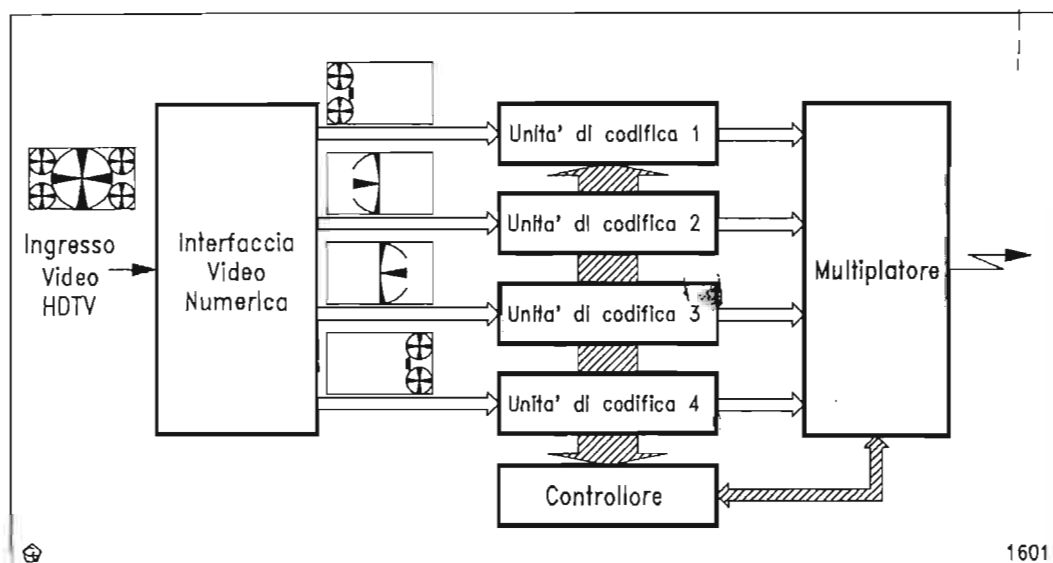



Fig. 2 — Il codificatore HDTV utilizzato nel corso delle trasmissioni Italia '90 è composto da una interfaccia d'ingresso che suddivide i campioni video attivi in quattro flussi. Ciascun flusso corrisponde ad uno dei quattro segmenti verticali in cui è suddivisa l'immagine HDTV. I dati risultanti sono memorizzati in un buffer e da qui prelevati, sotto forma di pacchetti, per essere multiplati e trasmessi. Il Controllore di buffer sovraintende alla organizzazione dei pacchetti, al loro prelievo ed alla determinazione del fattore di trasmissione in base all'occupazione dei quattro buffer.

formato HDTV europeo di tipo interlacciato (1250 righe per quadro e 25 quadri al secondo): se esso è campionato con una frequenza di campionamento pari a 54 MHz vengono generati 1440 campioni di luminanza e 720 campioni per ciascuna delle due componenti di crominanza in corrispondenza della porzione attiva di ogni riga di immagine.

Grazie alla interfaccia di ingresso del codificatore si creano quattro flussi pari a 720 campioni (di cui 360 di luminanza) per ciascuna delle 1152 righe attive costituenti i 25 quadri al secondo; ciascun flusso è dunque pari a 165,888 Mbit/s, essendo ciascun campione codificato con una parola di 8 bit. Questo flusso è esattamente quello corrispondente ad un segnale 4:2:2.

Il codificatore HDTV è quindi costituito da quattro unità di codifica identiche, e ciascuna di esse ha le capacità di elaborazione sufficienti per operare su un segnale 4:2:2. Questa struttura parallela ha consentito di utilizzare, anche nel caso del segnale HDTV, la tecnologia oggigiorno disponibile, con un alto livello di integrazione e operante ad un clock non superiore a 30 MHz. La struttura modulare permette di realizzare i complessi algoritmi richiesti per ridurre la quantità di dati da trasmettere di circa un fattore dieci: infatti in base ai valori precedentemente indicati il flusso binario in ingresso al codificatore è circa 660 Mbit/s, mentre la velocità di trasmissione sul canale via satellite è di circa 70 Mbit/s, comprensivi di audio e ridondanza per protezione dagli errori di canale (54 Mbit/s per il video).

Il codec così realizzato è in grado di operare con segnali HDTV conformi sia al formato europeo precedentemente citato, sia a quello proposto in Giappone e USA, cioè 1125 righe per quadro e 30 quadri al secondo. Nel corso della sperimentazione Italia '90 sono stati trasmessi e presentati al pubblico segnali prodotti secondo entrambi i formati. Il codec è inoltre in grado di operare con diverse velocità di trasmissione: nel caso specifico si è utilizzata una velocità pari a circa 70 Mbit/s in quanto questo valore costituisce il compromesso ottimale tra qualità del segnale decodificato e banda disponibile sul canale del satellite Olympus.

4. Prospettive per l'immediato futuro

Si è detto che l'attuale codec è in grado di elaborare segnali HDTV campionati a 54 MHz, con una definizione corrispondente a 1440 campioni di luminanza per la parte attiva della riga video. Tale precisione è largamente sufficiente se si considera la banda dei segnali attualmente forniti dalle telecamere HDTV e la definizione apprezzabile con i monitori e proiettori a grande schermo oggigiorno disponibili. Tuttavia i formati attualmente proposti per gli studi HDTV prevedono 1920 pel: per consentire la trasmissione del segnlae HDTV con qualità prossima a quella di studio e sufficiente per alcune operazioni di postproduzione, fra cui il «colour matte» o «intarsio», è in corso di realizzazione una nuova versione del codec in grado di elaborare la quantità di dati corrispondente anche agli ulteriori 960 campioni (480 di luminanza e 240 per ognuna delle crominanze).

Nel corso delle trasmissioni la qualità del segnale decodificato è risultata ottima, non distinguibile da quella del segnale all'uscita delle telecamere. Poiché nel caso di riprese sportive si fa ampio uso di carrellate o «panning», la stessa qualità è in genere ottenibile a velocità di trasmissione inferiori, in alcuni casi anche del 30%; questo ulteriore risparmio è ottenibile applicando tecniche di valutazione del movimento dal lato codificatore e di compensazione del movimento in co-decodifica. Tali tecniche permettono di ottenere una migliore predizione temporale e di conseguenza un maggior sfruttamento della ridondanza del segnale video. La trasmissione può quindi avvenire a pari qualità con velocità di segnalazione inferiore, oppure una migliore qualità è ottenibile a parità di velocità di trasmissione, garantendo così un margine nel caso di immagini particolarmente critiche e per quando, in futuro, saranno disponibili telecamere e dispositivi di visualizzazione capaci di una maggiore definizione.

Durante le trasmissioni di Italia '90 non sono stati riscontrati problemi dovuti agli errori introdotti del canale. Si è quindi dimostrano che il sistema di protezione utilizzato, basato sul codice per la correzione degli errori Bose-Chaudhuri-Hocquenghem BCH (255,239) è molto

Fig. 3 — Uno dei codificatori HDTV utilizzati per Italia '90. L'unità di conversione analogico/numerica, realizzata presso il Centro Ricerche RAI, fornisce il segnale numerico mediante un'interfaccia parallela su 16 bit più il clock al codificatore realizzato dalla Telettra. Il codificatore è racchiuso in un sottotelaio da 19" di larghezza e 6 unità di altezza. Il consumo è pari a circa 200 W. Le schede componenti il codificatore sono di 230 per 280 mm. Il flusso di dati in ingresso è suddiviso da una scheda di interfaccia nei quattro flussi corrispondenti ai segmenti verticali in cui è suddivisa l'immagine. Ciascun flusso è elaborato da due schede: la prima calcola i valori di predizione, sceglie il modo di codifica e calcola i coefficienti DCT; la seconda scheda codifica i coefficienti con parole VLC e le memorizza in un buffer di 4 Mbit. La scheda del controllore del buffer gestisce l'insieme dei quattro buffer appartenenti alle unità di codifica. Un'altra scheda provvede ad aggiungere la ridondanza per il codice per la protezione dagli errori di canale e a multiplare i dati provenienti dalle quattro unità di codifica video e da quella per la codifica dei segnali audio e dati ausiliari.

efficace. Ciononostante sarà presto disponibile, in alternativa al BCH, anche il codice Reed-Solomon RS (255,239), che offre prestazioni migliori in presenza di sequenze di errori (errori a burst). Tale codice è quello previsto dalla Raccomandazione 723.

Tutti questi miglioramenti, estensione a 1920 pel, compensazione del movimento e codice Reed Solomon, saranno introdotti nelle nuove versioni del codec sviluppate nella prima metà del 1991.

La valutazione delle prestazioni del sistema nel caso di immagini particolarmente critiche o in presenza di errori introdotti dal canale sono state oggetto di prove soggettive basate su simulazioni effettuate usando sequenze 4:2:2 (bibl. 7). I risultati di tali simulazioni sono state sostanzialmente confermati dalle prove effettuate utilizzando il codec reale, sia nella versione atta ad elaborare i segnali 4:2:2, sia quelli HDTV. Ulteriori prove soggettive sono già in programma, per completare la valutazione del sistema attuale e per consentire lo studio di ulteriori miglioramenti ed estensioni.

Le attività di ricerca e di ottimizzazione relative alle tecniche di riduzione della ridondanza del segnale video e di adattamento delle stesse alle varie applicazioni, ad esempio alla trasmissione a velocità di segnalazione costante oppure variabile (bibl. 8) ed alla registrazione, non accennano a diminuire, come testimonia la costituzione di diversi gruppi di normalizzazione e l'avvio di nuovi progetti di ricerca.

BIBLIOGRAFIA

1. - Barbero M., Del Pero R., Muratori M., Stroppiana M.: *Bit-Rate Reduction Techniques based on DCT for HDTV Transmission*. SUPERCOMM ICC '90, Atlanta USA, Aprile 1990.

2. - Ardito M., Barbero M.: *Point-to-Multipoint Transmission of TV and HDTV Signals*. ITU/CyBC Seminar on New Technologies in Sound and Television Broadcasting, Nicosia Cipro, Ottobre 1990.

3. - Rao K. R., Yip P.: *Discrete Cosine Transform: Algorithms, Advantages, Applications.*, Academic Press Inc., 1990

4. - Bellora R., Dimino G., Muratori M.: *Hybrid DCT: comparison of the statistics of DCT coefficients and processing modes with and without motion compensation*, in «Signal Processing of HDTV, 2», Elsevier Science Publishers (Proc. of the 3rd Int. Workshop on HDTV, Torino I, Agosto 1989)

5. - Del Pero R., Giromini P., Morello A., Barbero M.: *Criteria for the protection of the video information in a codec based on DCT*, in «Signal Processing of HDTV, 2» Elsevier Science Publishers (Proc. of the 3rd Int. Workshop on HDTV, Torino I, Agosto 1989.

6. - Barbero M., Cucchi S., Hernando Bailon J. L.: *A Flexible Architecture for a HDTV Codec based on DCT*, in «Signal Processing of HDTV, 2», Elsevier Science Publishers (Proc. of the 3rd Int. Workshop on HDTV, Torino I, Agosto 1989.

7. - Ardito M., Barbero M., Ibañez D.: *Performance of Codecs for Bit-Rate Reduction applied on Conventional TV and HDTV Signals*, IBC '90, Brighton UK, Settembre 1990.

8. - Barbero M., Del Pero R., Ravera S., Rossato L.: *The use of bitrate reduction techniques for the transmission of video signals*, DSP90 - 2nd Int. Workshop on Digital Signal Processing Techniques Applied to Space Communications, Torino I, Settembre 1990.

ITALIA '90: STAZIONI TRASMITTENTI DI SEGNALI HDTV NUMERICI VIA SATELLITE OLYMPUS

F. BONACOSSA, G. MORO, B. SACCO, D. TABONE*

SOMMARIO — Per trasmettere in HDTV numerica le partite della recente Coppa del Mondo di Calcio, la RAI ha impiegato due stazioni trasmittenti verso il satellite Olympus. Per le partite di Roma e per le repliche si è adattata alla trasmissione numerica la stazione fissa di Roma RAI CPTV, normalmente impiegata come stazione di salita di RAI-SAT. Per le partite giocate negli altri stadi la RAI ha impiegato una stazione trasportabile espressamente progettata per segnali numerici. In questo articolo si illustrano le caratteristiche tecniche e la funzionalità delle due stazioni.

SUMMARY — ***ITALIA '90: Up-link earth stations for digital HDTV signals via Olympus satellite*** *- As regards the digital HDTV transmission of the matches played at the FIFA World Cup 1990 in Italy, the RAI used two up-link earth stations to the Olympus satellite. The fixed up-link earth station of the RAI Production Centre in Rome, generally used as the feeder-link of RAI-SAT, was adapted to transmit the matches played in Rome and the repeats. As for the matches played in the other stadiums the RAI used a transportable earth station designed for digital signals on purpose. This paper illustrates the technical characteristics and the operational capability of the two stations.*

1. Introduzione

In altro articolo, su questo stesso numero della rivista, si descrive l'organizzazione generale dell'esperimento di HDTV digitale condotto dalla RAI durante la recente Coppa del Mondo di Calcio, Italia '90.

L'obiettivo era di mostrare in alcuni punti di visione, predisposti in Italia a cura della RAI, la qualità delle immagini delle partite riprese in alta definizione e trasmesse via satellite con codifica e modulazione digitale.

Le sedici partite oggetto dell'esperimento in alta definizione numerica sono state trasmesse via satellite Olympus: con la stazione trasmittente fissa di Roma, presso il Centro di produzione TV della RAI, sono state trasmesse le partite giocate allo stadio Olimpico di Roma, mentre quelle giocate negli stadi di Torino, Milano, Firenze, Napoli sono state trasmesse mediante una stazione trasportabile costruita appositamente dalla Selenia Spazio su specifiche fornite dal *Centro Ricerche* RAI (vedi bibl. 1).

Altri articoli illustrano come le partite di Roma siano state riprese secondo lo standard di produzione europeo (1250/50/2:1) con telecamere sviluppate nell'ambito del progetto Eureka Eu-95, mentre le altre partite sono state oggetto di una coproduzione RAI-NHK con riprese secondo lo standard giapponese (1125/60/2:1) con l'impiego di mezzi itineranti NHK affiancati alla stazione mobile RAI (Fig. 1).

In questo articolo si esaminano le caratteristiche delle due stazioni impiegate durante l'esperimento.

2. Uso delle stazioni trasmittenti trasportabili

I vantaggi importanti che le comunicazioni via satellite offrono ai radio-diffusori sono noti da tempo.

In tema di riprese esterne l'uso del satellite consente di «esercire» qualunque località del territorio nazionale pur in presenza di difficoltà legate alla montuosità del posto o, come nelle città, alla vicinanza di edifici di ostacolo al collegamento.

Una stazione trasportabile in grado di trasmettere al satellite permette di inviare i segnali del programma ver-

Fig. 1 — La stazione trasportabile, costruita dalla Selenia Spazio, in postazione presso lo Stadio Olimpico di Roma.

* P.i. Fulvio Bonacossa, ing. Giovanni Moro, ing. Bruno Sacco, p.i. Dario Tabone del Centro Ricerche RAI - Torino.
Dattiloscritto pervenuto alla redazione il 30 ottobre 1990.

so il centro di Controllo, o, come nel caso di Italia '90, direttamente verso i punti visione quando si tratti di esperimento punto-multipunto.

3. Il segmento spaziale

È stato usato il satellite Olympus e non il già da tempo usato ECS, perché calcoli preliminari e poi la sperimentazione notturna con il canale RAI di ECS, fatta dalla stazione del Fucino, con la fattiva collaborazione di Telespazio (vedi bibl. 2) cui va il nostro ringraziamento doveroso, avevano chiaramente dimostrato che non era tecnicamente possibile ottenere la qualità indispensabile e margini adeguati quando il segnale dagli stadi fosse stato trasmesso con le potenze fornite dalle stazioni ECS della serie ITA, in dotazione al pool RAI-TELESPAZIO, indipendentemente dalle dimensioni delle antenne riceventi.

Il satellite Olympus invece presenta in ricezione un G/T migliore di 4 dB rispetto all'ECS; era inoltre possibile progettare una stazione trasportabile con EIRP adeguata e l'esempio era fornito dalla TDS-5 costruita dalla Selenia Spazio, su ordine ESA, per esprerimenti e dimostrazioni nel pacco di diffusione diretta a 12 GHz di Olympus (canale 20 a copertura europea), con segnali analogici secondo lo standard MAC.

Il transponder del canale 24 del satellite Olympus, usato per la sperimentazione, ha una frequenza in salita di 17,768 GHz, una frequenza di discesa di 12,168 GHz, un fattore G/T = 6 dB/K e una EIRP di 63,3 dBW. per una descrizione più dettagliata si rimanda al numero speciale di questa rivista sul satellite Olympus.

4. Stazioni mobili

Le peculiarità di una stazione mobile sono tutte riconducibili all'esigenza di contenere peso e dimensioni del mezzo al fine di favorirne le qualità operative e funzionali.

L'antenna, per esempio, se fatta di un disco parabolico intero, deve avere un diametro limitato dalla soluzione adottata per la sua posizione in condizione di trasporto, e sarà comunque compatibile con la sagoma di ingombro del mezzo, stabilita dalla regolamentazione dei trasporti su strada.

Anche le soluzioni possibili al problema dell'alimentazione della stazione saranno condizionate alle esigenze di maneggevolezza del mezzo, infatti la necessità di installare tubi di trasmissione di elevata potenza per compensare la relativa piccolezza dell'antenna, molto spesso non consente di progettare un mezzo autosufficiente in termini di alimentazione di energia elettrica.

Nel caso della stazione mobile RAI per Olympus, queste esigenze hanno portato ad una realizzazione con le seguenti caratteristiche:

— la stazione è integrata con l'automezzo IVECO 145/17;
— tutti gli apparati sono alloggiati in unico locale con l'eccezione degli amplificatori di ricezione a basso rumore;
— la lunghezza totale del mezzo contenuta in 9,1 metri;
- l'antenna ha diametro 3,3 metri con illuminazione tipo Cassegrain ed è montata su una piattaforma girevole che consente un movimento di azimuth di ±110°;
— non ci sono parti da smontare;
— l'energia elettrica 380 Volt/50 Hz trifase deve essere fornita dall'esterno, rete o complesso di continuità.

La stazione mobile della RAI è la prima stazione conosciuta al mondo, concepita per trasmettere i segnali numerici a 18 GHz. Può lavorare in Europa e nel Nord Africa da tutte le località entro la regione coperta dall'antenna ricevente del satellite (vedi fig. 5 della bibl. 3), è in grado di trasmettere ad entrambi i canali di diffusione diretta di Olympus e, per quanto fattibile, anche ad altri satelliti per DBS funzionanti secondo il Piano WARC '77; inoltre può ricevere qualunque canale DBS del Piano su entrambi i sensi di polarizzazione.

Si è già accennato al fatto che la stazione trasportabile è stata studiata e progettata per l'utilizzazione nella trasmissione di segnali numerici, mentre la stazione fissa, inizialmente concepita per la trasmissione di segnali TV/FM, è stata adattata in modo che le prestazioni finali consentissero condizioni di trasmissioni compatibili con le specifiche del modulatore-demodulatore a 70 Mbit/s di cui la stazione doveva essere equipaggiata.

Vediamo quindi il funzionamento del modem impiegato.

5. Il modem

Come già accennato per la trasmissione del segnale HDTV via Olympus è stata adottata la modulazione QPSK, ad un bit rate di 70 Mbit/s (69,870 Mbit/s, per la precisione).

Lo spettro del segnale modulato ha una banda a 3 dB di circa 40 MHz; la catena stazione trasmittente — satellite — stazione ricevente deve comunque avere una banda passante superiore (anche in previsione dell'uso per servizi diversi).

Il canale di trasmissione si può considerare non distorcente se la risposta d'ampiezza e il ritardo di gruppo sono costanti su una banda che, per il caso del QPSK, si può assumere larga come metà della frequenza associata al bit rate trasmesso. In particolare la piattezza in banda può essere quantificata in qualche decimo di decibel, e il ripple sul ritardo di gruppo in una frazione del tempo di simbolo.

Il modulatore riceve in ingresso due flussi di dati in formato HDB3, codifica di linea che permette di inglobare i dati e il clock, in un unico segnale. Ciascuno dei due flussi in ingresso ha un bit rate di 34,368 Mbit/s, compatibile con le reti numeriche in gerarchia e con evidenti semplificazioni nel trasporto terrestre del segnale digitale. All'interno del modulatore i due flussi vengono multiplati in una trama da un circuito con memoria elastica utilizzato anche per compensare il jitter della temporizzazione dei segnali in arrivo, spesso presente dopo il transito su reti di terra.

La frequenza della portante è 140 MHz; il filtraggio è del tipo coseno rialzato con roll-off 50% ripartito egualmente tra trasmissione e ricezione.

Il circuito di recupero della portante, utilizzata per la demodulazione coerente, ha un campo di aggancio di ± 1,1 MHz, valore che impone un vincolo sulla tolleranza di frequenza del segnale ricevuto, e quindi sulla stabilità degli oscillatori di conversione di tutta la catena trasmissiva.

L'uscita del demodulatore fornisce due segnali a 34 Mbit/s con formato HDB3.

6. Adattamento delle stazioni di terra alle trasmissioni numeriche

Per le due stazioni si sono quindi potuti applicare criteri comuni di progetto e di intervento che si possono riassumere nei punti seguenti:

— ingresso e uscita IF delle stazioni a 140 MHz per consentire l'impiego del modem citato;

— banda passante e ritardo di gruppo delle catene di trasmissione e ricezione (IF-RF-IF) devono rispondere alle seguenti specifiche:

banda passante: 80 MHz

risposta di ampiezza	entro ± 0,4 dB nella banda Fo ± 20 MHz
	entro ± 0,7 dB nella banda Fo ± 36 MHz
ritardo di gruppo	entro 4 ns_{pp} nella banda Fo ± 18 MHz
	entro 8 ns_{pp} nella banda Fo ± 36 MHz

— alta stabilità delle conversioni di frequenza; per il rumore di fase si assumono le seguenti specifiche:

< – 60 dBc/Hz a 500 HZ < – 90 dBc/Hz a 100 kHz

— particolare attenzione è stata posta nell'assicurare un funzionamento in zona lineare della catena trasmittente sino all'ingresso del tubo amplificatore di potenza finale in modo da rendere la dispersione dello spettro del segnale irradiato dipendente solo dal punto di lavoro del klystron, controllando la situazione di interferenza nei canali di salita adiacenti;

— possibilità di inserire un filtro di canale a frequenza centrale 17,768 GHz tra l'amplificatore finale e l'antenna per limitare le interferenze sui canali adiacenti, in particolare sui canali 22 e 26, occupati sulla up-link dai segnali del satellite tedesco TV-SAT. La risposta del filtro deve risultare:

1 dB a ± 20 MHz; 3 dB a ± 25 MHz; 20 dB a ± 38 MHz

Dalle prove preliminari il filtro di uscita, nel caso della stazione di Roma, è risultato inutile in quanto la potenza in salita necessaria per garantire i margini richiesti per l'intero collegamento era ottenuta con l'amplificatore di potenza funzionante in zona lineare. Al contrario, nel caso della stazione mobile, il filtro di uscita risultava opportuno in considerazione della minore EIRP, o potenza irradiata, disponibile, e per questa ragione è stato inserito con possibilità di esclusione;

— equipaggiamento aggiuntivo con strumenti di monitoraggio della qualità del collegamento numerico.

7. Specifiche della stazione mobile:

Banda di frequenza di trasmissione		17,3 – 18,1	GHz
Banda di frequenza di ricezione		11,7 – 12,5	GHz
Frequenza intermedia		70 e 140	GHz
Presintonie		canali 20, 24, 28 WARC '77	
Stabilità di frequenza		10^{-7}/mese	
Frequenza di traslazione del TLT		5600	MHz
EIRP portante singola (modo A)		85	dBW
EIRP di ciascuna delle due portanti (modo B)		79	dBW
Spurie fuori canali		< 10	dBW/MHz
G/T a 11,7 GHz e cielo chiaro		25	dB/K
Polarizzazione TX		circolare doppia	
Polarizzazione RX		circolare doppia	
Isolamento tra le polarizzazioni		30	dB
Diagramma	$1° < \phi < 36°$	$29 - 25 \log\phi$	dBi
di irradiazione TX	$\phi > 36°$	– 10	dBi
Diagramma	$1° < \phi < 36°$	$32 - 25 \log\phi$	dBi
di irradiazione RX	$\phi > 36°$	– 7	dBi
Movimento EL		0° ÷ 60°	
Movimento AZ		± 110°	
Larghezza di banda	senza filtro	80	MHz
	con filtro	50	MHz
Conversione AM/PM	catena TX	< 5	°/dB

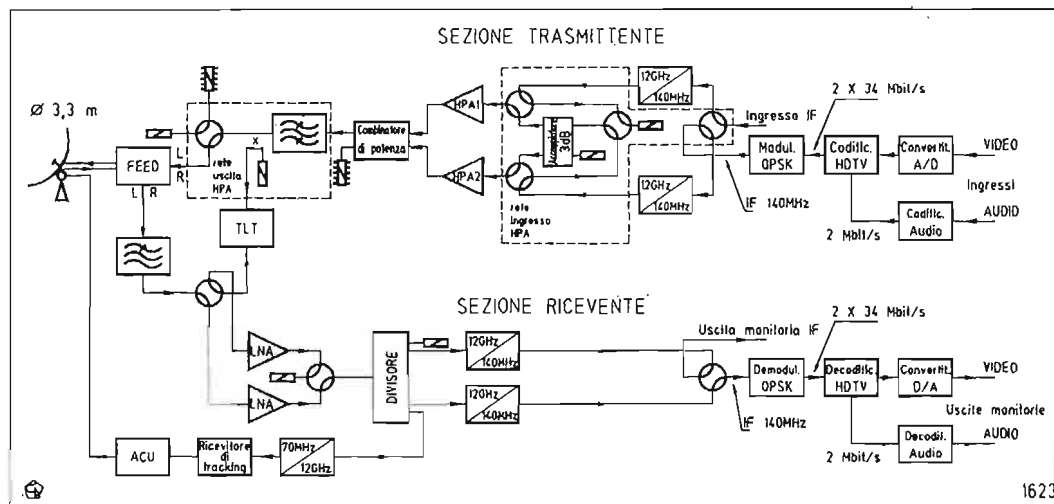



Fig. 2 — Schema a blocchi della stazione trasportabile.

8. Schema a blocchi della stazione mobile

In figura 2 è riportato lo schema a blocchi semplificato della stazione mobile.

Gli ingressi analogici video e audio, dopo opportuna equalizzazione, vengono inviati ai rispettivi convertitori A/D e quindi alla catena di codifica e modulazione numerica.

Il segnale in uscita dal modulatore QPSK a frequenza 140 MHz può essere, tramite opportuno deviatore nella rete di ingresso HPA, instradato verso l'uno o l'altro dei convertitori in salita 140 MHz/12 GHz a sintesi di frequenza. La catena di trasmissione già a questo punto mostra la ridondanza 1 + 1. Le uscite del convertitore di frequenza possono essere alternativamente instradate all'ingresso di un accoppiatore 3 dB, le cui due uscite alimentano i due amplificatori di alta potenza (HPA) di cui la stazione è fornita. La regolazione dei livelli nei singoli stadi della catena di trasmissione sino a questo punto deve essere tale da garantire un funzionamento in zona lineare. I due amplificatori di alta potenza HPA 1 e HPA 2, ciascuno con potenza massima di uscita pari a 1350 watt, possono essere in uscita combinati in fase per dare il massimo di EIRP disponibile su portante singola (questa configurazione riportata nel paragrafo precedente è descritta come modo A). In questa configurazione i due tubi a klystron sono posti in parallelo e ciò si ottiene agendo su di un regolatore di fase inserito all'ingresso di uno dei due amplificatori di potenza e non riportato per semplicità nello schema a blocchi. In caso di perdita di un HPA il combinatore di potenza è automaticamente by-passato e la perdita di potenza netta è di soli 3 dB (protezione *soft-fail*).

Un'altra configurazione è possibile, descritta come modo B: essa consente la trasmissione contemporanea di due portanti separate, ciascuna con potenza di circa 600 watt, ma con perdita della ridondanza. Infatti, i due upconverter 140 MHz/12 GHz sintonizzati su frequenze diverse sono separatamente alimentati da due segnali di frequenza intermedia e la rete di ingresso HPA consente di pilotare separatamente i due amplificatori di potenza. L'uscita del combinatore di potenza può essere chiusa su una terminazione per realizzare una situazione di stand-by, oppure può essere instradata verso l'antenna nella condizione di trasmissione.

Sul percorso in guida d'onda dal combinatore di potenza all'antenna è inseribile il filtro di canale a 18 GHz già descritto.

Tramite due accoppiatori direzionali in guida d'onda sono realizzati punti di prelievo del segnale per fornire sia una monitoria di potenza, sia il segnale di ingresso al così detto circuito TLT di prova.

La funzione del TLT è sinteticamente quella di convertire di frequenza i 18 GHz di uscita nel corrispondente canale nella banda di ricezione a 12 GHz e di attenuare il livello del segnale in modo da renderlo compatibile con l'ingresso di uno degli amplificatori a basso rumore di ricezione al fine di permettere un controllo della catena di trasmissione e di ricezione, esclusa l'antenna e il suo feed.

Proseguendo nella descrizione del circuito verso l'antenna lo schema a blocchi mostra simbolicamente la possibilità di trasmettere indifferentemente con polarizzazione circolare oraria o antioraria. L'antenna è una parabola del diametro di 3,3 metri in lastra di alluminio, illuminata con geometria tipo Cassegrain, i cui movimenti di elevazione e di azimuth sono controllabili da comandi manuali o asserviti alla ricezione del segnale di beacon trasmesso dal satellite Olympus a frequenza 12501,866 MHz. La stessa antenna riceve i segnali di ritorno dal satellite indifferentemente con la polarizzazione circolare destra o sinistra. Il segnale ricevuto di livello molto basso transita in un filtro rigetta-banda centrato alla frequenza di trasmissione; questo per impedire che la catena di ricezione venga saturata da un segnale ad alto livello. Il segnale ricevuto è presentato all'ingresso di uno dei due amplificatori a basso rumore in ridondanza 1 + 1. Successivamente un divisore di potenza invia il segnale a 12 GHz all'ingresso dei due convertitori di frequenza in discesa 12 GHz-140 MHz. Il segnale convertito alimenta la catena di demodulazione QPSK e di decodifica HDTV. Una adeguata strumentazione monitoria permette di valutare le prestazioni globali del sistema.

9. Specifiche della stazione fissa

Banda di frequenza di trasmissione		17,3 – 18,1	GHz
Banda di frequenza di ricezione		11,7 – 12,5	GHz
Frequenza intermedia		70	GHz
Presintonie		canali 20, 24, 28 WARC '77	
Stabilità di frequenza		± 10 kHz	
Frequenza di traslazione del TLT		5600	MHz
EIRP portante singola (modo A)		86	dBW
EIRP di ciascuna delle due portanti (modo B)		86	dBW
Spurie fuori canali		< 4	dBW/4 MHz
G/T a 11,7 GHz e cielo chiaro		32	dB/K
Polarizzazione TX		circolare doppia	
Polarizzazione RX		circolare doppia	
Isolamento tra le polarizzazioni		30	dB
Diagramma	$1° < \phi < 36°$	$29 - 25 \log\phi$	dBi
di irradiazione TX	$\phi > 36°$	– 10	dBi
Diagramma	$1° < \phi < 36°$	$29 - 25 \log\phi$	dBi
di irradiazione RX	$\phi > 36°$	– 10	dBi
Movimento EL		5° ÷ 65°	
Movimento AZ		± 60°	
Larghezza di banda	senza filtro	80	MHz
	con filtro	50	MHz
Conversione AM/PM	catena TX	< 5	°/dB

10. Schema a blocchi della stazione fissa

In figura 3 è riportato lo schema a blocchi semplificato della stazione fissa di Roma.

Come si può osservare, la stazione era composta da due parti interconnesse tra loro mediante collegamenti in fibra ottica: una, di generazione vera e propria, situata nei pressi dello Stadio Olimpico; l'altra, con gli apparati di trasmissione, ricezione e controllo, all'interno del Centro di Produzione TV di via Teulada.

10.1 *Postazione Stadio Olimpico*

I segnali audio (3 canali: destro, sinistro e effetti) e video (Y, P_R, P_B e sincronismi a tre livelli) analogici di in-

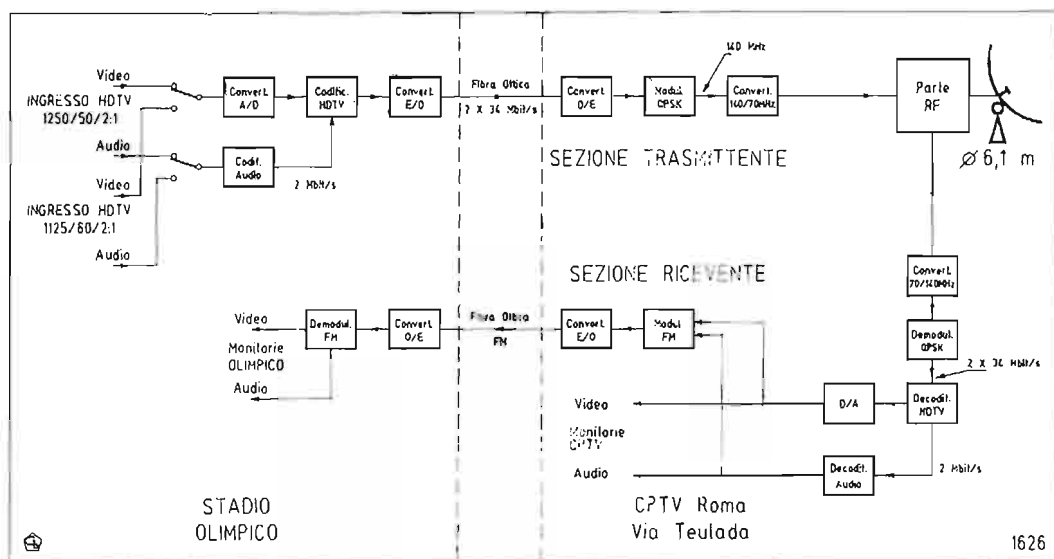



Fig. 3 — Schema a blocchi della stazione fissa.

gresso relativi alle sorgenti con standard HDTV europeo o giapponese, dopo opportuna equalizzazione, venivano inviati rispettivamente al codificatore numerico a 2 Mbit/s e al convertitore A/D che alimentava il codificatore HDTV vero e proprio.

I due flussi a 34 Mbit/s, mediante collegamento in fibra ottica, erano inviati al CPTV RAI di via Teulada a Roma.

Il corretto inserimento dei segnali era controllabile mediante opportuna strumentazione audio/video, oltre che soggettivamente con monitor HDTV e casse acustiche. Mediante fibre ottiche modulate in frequenza, provenienti dalla stazione fissa del CPTV, era altresì possibile il controllo, alla postazione dello Stadio Olimpico, del segnale irradiato e ricevuto dal satellite Olympus con l'antenna del CPTV.

10.2 *Stazione fissa*

Come detto, la stazione fissa è la stessa utilizzata per le trasmissioni sperimentali in PAL di RAI-SAT, in cui al posto del modulatore MF era stato inserito il modulatore numerico a 70 Mbit/s e FI = 140 MHz. Poiché la stazione è stata costruita per funzionare con una FI a 70 MHz, è stato necessario interporre un convertitore di frequenza 140/70 MHz.

Mediante gli apparati della stazione di ricezione veniva alimentata la parte di controllo: il segnale FI a 70 MHz, dopo conversione a 140 MHz, entrava nel demodulatore a 70 Mbit/s che forniva in uscita i due flussi a 34 Mbit/s per il decodificatore HDTV.

Il convertitore D/A e il decodificatore audio a 2 Mbit/s fornivano i segnali analogici video e audio che erano controllabili in locale ed inviati alla sala controllo appositamente preparata in un locale del CPTV stesso.

Nello schema a blocchi di figura 3 la parte a radiofrequenza della stazione di salita di Roma è rappresentata da un semplice blocco perché, come già detto, le stazioni di terra hanno architetture molto simili tra loro, per cui anche per questa stazione varrebbe l'illustrazione fatta al punto 8 con alcune differenze specifiche che qui riportiamo.

I due amplificatori di alta potenza (1350 watt ciascuno come sulla stazione trasportabile) non sono combinabili in fase, ma essi operano nella configurazione ridondata, designata come modo A, il primo in linea ed il secondo in *stand-by*, oppure sommati e su frequenze distinte per la trasmissione simultanea di due portanti TV nella configurazione non ridondata, designata come modo B.

Fig. 4 — Apparati di codifica e controllo del segnale HDTV numerico presso lo Stadio Olimpico.

Stazione trasmittente fissa di Telespazio puntata verso il satellite Olympus e posta sull'edificio del Centro di Produzione TV della RAI a Roma. L'antenna ha un diametro di 6,1 metri, con illuminatore tipo Cassegrain, ed è munita di sistema automatico di puntamento del satellite.
La stazione, trasmittente nella banda di frequenze di 17 GHz, ha inoltre capacità di ricezione monitoria nella banda a 12 GHz. La stazione è stata costruita dalla Società SELENIA SPAZIO.

Anche in assenza di combinazione in fase, la EIRP di stazione risultante è pari a 86 dBW perché l'antenna di diametro 6,1 metri ha un guadagno in trasmissione pari a 59,6 dBi.

Analogamente alla stazione trasportabile anche per l'antenna della stazione fissa la geometria di illuminazione è di tipo Cassegrain.

Per quanto riguarda la ricezione nella stazione è installata un'unica catena monitoria.

11. Conclusioni e ringraziamenti

Si sono descritte le caratteristiche delle due stazioni di terra, fissa e mobile, operanti con il pacco di diffusione diretta del satellite Olympus, impiegate dalla RAI per il recente esperimento di trasmissione punto-multipunto di HDTV numerica.

Le stazioni hanno operato in maniera soddisfacente durante l'esperimento e gli autori non possono non ricordare la preziosa collaborazione prestata prima e durante la manifestazione degli Ingg. Francesco Fiorica, Carlo Scaffidi e Stefano Fioravanti della Società Selenia Spazio, costruttrice delle stazioni di terra, nonché dall'Ing. Massimo Pronti e dal Sig. Giuseppe di Mattia della Società Telespazio. (3931)

BIBLIOGRAFIA

1. - G. P. Pacini: *Stazione mobile per trasmissione a Olympus. Pacco 18-12 GHz*, «RAI-Centro Ricerche», Relazione Tecnica n. 88/14/I.

2. - G. Moro, G. Garazzino, M. Tabone: *Trasmissioni sperimentali QPSK a 34 e 70 Mbit/s via Eutelsat I*, «RAI-Centro Ricerche», Relazione Tecnica n. 90/10/I.

3. - G.P. Pacini: *Satellite Olympus-F1*, «Elettronica e Telecomunicazioni», n. 2/3, 1989.

ITALIA '90: POSTAZIONI RICEVENTI PER TRASMISSIONI PUNTO-MULTIPUNTO DI HDTV NUMERICA DA SATELLITE

M. ARIAUDO, G. CERRUTI, G. GARAZZINO*

SOMMARIO — Si descrivono le postazioni riceventi allestite dalla RAI in occasione dei Campionati Mondiali di Calcio «Italia '90» per la sperimentazione di HDTV numerica via satellite Olympus, con due diversi standard (1250 righe/50 Hz e 1125 righe/60 Hz).
Ogni postazione comprende una sezione a radiofrequenza (l'antenna parabolica e le due unità di conversione), attraverso la quale il segnale, modulato QPSK con un flusso a 70 Mbit/s, è traslato dalla frequenza di ricezione del canale 24 del satellite (12168, 62 MHz) alla frequenza intermedia di 140 MHz ed inviato al demodulatore. Da questo sono estratti due flussi a 34 Mbit/s dai quali il decodificatore HDTV ricostruisce il segnale video numerico e il flusso a 2 Mbit/s dei 3 canali audio numerici. Attraverso il decodificatore audio si ottengono i segnali audio analogici per gli altoparlanti. Analogamente, per il segnale video numerico si opera la conversione D/A e, per mezzo di due differenti matrici (50 Hz e 60 Hz), si ottengono i segnali R, G, B per il proiettore HDTV. A causa dell'ampia gamma di frequenze e delle speciali esigenze imposte dal segnale numerico, gli apparati costituenti le postazioni riceventi sono di notevole complessità e richiedono caratteristiche molto più severe di quelle dei ricevitori MF da satellite esistenti sul mercato. Perciò, per la maggior parte degli apparati (unità interna, decodificatore numerico, generatori di sincronismi, matrici, equalizzatori/distributori video), è stato necessario effettuarne la progettazione presso il Centro Ricerce RAI e coordinarne la realizzazione presso l'industria nazionale. È stato altresì necessario modificare i monitori HDTV di controllo per renderli adatti al funzionamento bistandard. Per gli apparati acquisiti all'esterno, si è invece resa necessaria una accurata selezione, con sofisticate misure volte a quantificare caratteristiche non specificate (stabilità di frequenza e rumore di fase, particolarmente importanti per i segnali numerici). Negli oltre 40 giorni di attività durante ***ITALIA '90****, il funzionamento delle otto stazioni riceventi, dislocate sul territorio nazionale, è risultato assolutamente soddisfacente ed affidabile.*

SUMMARY — ***Receiving stations for point-to-multipoint transmission of digital HDTV signals via satellite:*** *- The article describes the receiving stations properly equipped by RAI on the occasion of the FIFA World Cup 1990 (Italia '90) for the experiment of digital HDTV via the Olympus satellite, adopting two different formats (1250 lines/50 Hz and 1125 lines/60 Hz). Each station was equipped with a radiofrequency section (parabolic antenna and two conversion units), through which the QPSK-modulated signal at 70 Mbit/s is translated from the satellite channel 24 reception frequency (12168.62 MHz) to the 140 MHz intermediate frequency, and is sent to the demodulator. From the demodulator two 34 Mbit/s streams are extracted from which the HDTV decoder reconstructs the digital video signal and the 2 Mbit/s stream of the three digital sound channels. The analogue sound signals are obtained with the audio decoder and sent to the loudspeakers. Analogously, the D/A conversion is made for the digital video signal and, by means of two different matrices (50 Hz and 60 Hz) the signals R, G, B are obtained for the HDTV projector.*
Owing to the wide frequency range and to the particular requirements asked by the digital signal, the units of the receiving stations are very complex and require more severe characteristics than those of the satellite FM receivers currently on the market. Hence, for most of the equipment (internal unit, digital decoder, sync generators, matrices, video equalizers/distributors) it was necessary to make its design at the RAI Research Centre and to co-ordinate its implementation with the national industry. It was also necessary to modify the HDTV monitors so as to make them suitable for operating according to the two formats. Conversely, as regards the units bought outside an accurate selection was required, adopting sophisticated tests to quantify not specified characteristics (frequency stability and phase noise, particularly important for the digital signals). During more than 40 days at ***ITALIA '90*** *the operation of the eight receiving stations located over the national territory resulted to be completely satisfactory and reliable.*

1. Introduzione

Gran parte dell'impegno realizzativo inerente alla messa in opera dell'intero complesso di trasmissione e ricezione di segnali video numerici HDTV, realizzato dalla RAI in occasione dei Campionati Mondiali di Calcio del 1990 (bibl. 1), si è concentrato nelle postazioni riceventi.

Ciò è facilmente comprensibile se si considera che le

(*) Dott.sa Margherita Ariaudo, p.i. Giovanni Cerruti, p.i. Giorgio Garazzino del Centro Ricerche RAI - Torino.
Dattiloscritto pervenuto alla redazione il 18 ottobre 1990.

postazioni dovevano essere almeno otto ed i tempi realizzativi non erano superiori ai quattro mesi, entro i quali si doveva provvedere non solo all'assemblaggio degli apparati necessari, ma anche alla realizzazione degli apparati non ancora disponibili sul mercato.

Gli apparati di ricezione progettati e realizzati per l'occasione dal *Centro Ricerche* della *RAI* sono:

— per la sezione a radio frequenza, le unità interne;
— per il segnale video numerico, il convertitore D/A;
— per il segnale video di banda base, i generatori di sincronismi (PG), le matrici e, in parte, gli equalizzatori/distributori;
— gli apparati audio ausiliari.

Una notevole complicazione derivava dall'esigenza di poter lavorare indifferentemente sia a standard 50 Hz che 60 Hz. Questa esigenza ha inoltre richiesto modifiche circuitali dei monitori HDTV di controllo (monitori Sony), in quanto non disponibili con funzionamento bistandard.

La vasta gamma di frequenze dei segnali da trattare (dai 12 GHz della radiofrequenza ricevuta dal satellite, fino alla banda base del segnale video) e la presenza di segnali video sia numerici che analogici, ha comportato il coinvolgimento di una equipe assai eterogenea di specialisti.

Si è dovuto poi curare il coordinamento delle ditte esterne alle quali era demandata la realizzazione di vari apparati, frutto di reciproca collaborazione: particolarmente importante la collaborazione con la ditta Telettra, realizzatrice di uno degli apparati più complessi, il codecodificatore numerico HDTV, apparato la cui fattibilità era già stata dimostrata per simulazione presso il Centro Ricerche Rai, ma la cui complessità realizzativa richiedeva l'opera di una industria elettronica ai massimi livelli tecnologici.

2. Descrizione a blocchi della postazione ricevente

La postazione ricevente, come mostra lo schema a blocchi di figura 1, si articola in tre parti:

— l'antenna e l'unità esterna, alloggiate in luogo adatto alla ricezione del segnale del satellite;

— l'unità interna, il demodulatore QPSK, il decodificatore video numerico, il convertitore video D/A, il decodificatore audio numerico, le matrici 50 e 60 Hz, gli equalizzatori/distributori video, i PG a 50 e 60 Hz, i distributori audio, i pannelli di commutazione, il monitore ed il waveform monitor di controllo. Tutti questi apparati sono alloggiati su due rack posti in una saletta di regia;

— un apparato di proiezione video, i diffusori audio destro e sinistro e la catena di diffusori audio per l'effetto surround, posti nella sala visione.

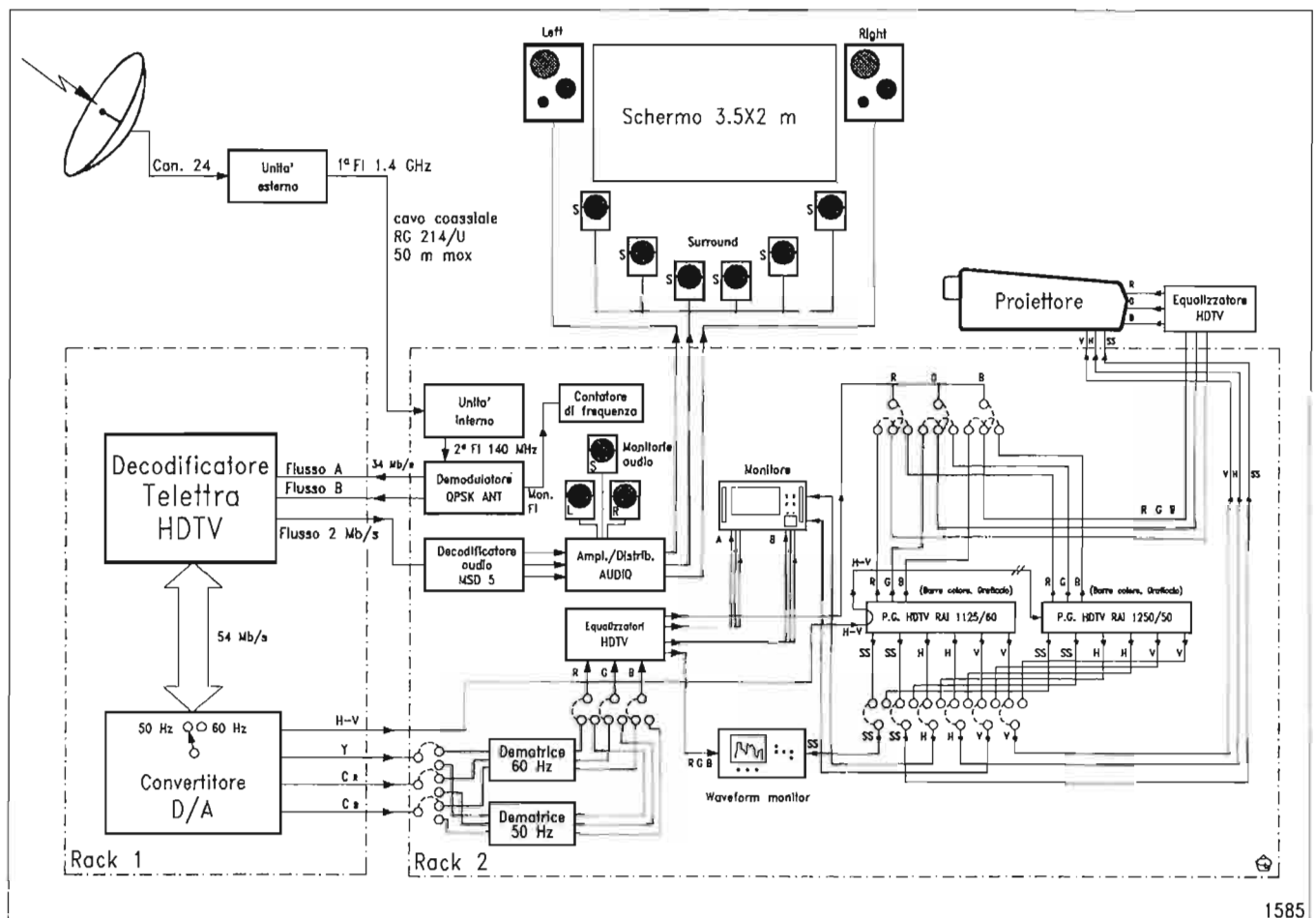



Fig. 1 — Schema a blocchi della postazione ricevente.

Il segnale, modulato QPSK a 70 Mbit/s, più precisamente a 69,87 Mbit/s, è irradiato dal satellite alla frequenza di 12168,62 MHz (canale 24). A terra viene ricevuto dall'antenna e immediatamente convertito ad una prima FI di 1418,62 MHz per mezzo dell'unità esterna. Tale FI è più adatta ad alimentare il cavo coassiale di collegamento alla unità ricevente posta all'interno dell'edificio, collegamento che normalmente risulta di lunghezza non indifferente (molte decine di metri).

Il suddetto segnale a RF raggiunge la postazione ricevente ed entra nella unità interna, il cui compito è quello di convertire la prima FI di 1418,62 MHz in una seconda FI di 140 MHz, con caratteristiche di notevole stabilità di frequenza: compito non facile in considerazione del divario tra le due FI e della scarsa dinamica della frequenza di ingresso accettata dal successivo demodulatore QPSK.

Il processo di demodulazione è in questo caso particolarmente delicato, a causa del rumore e delle distorsioni sul collegamento.

Il demodulatore QPSK riceve il segnale alla seconda FI ed estrae due flussi di segnali numerici su linee coassiali parallele, ciascuno a 34 Mbit/s, più precisamente a 34,368 Mbit/s, coerentemente con la filosofia operativa del decodificatore video numerico RAI/Telettra.

Il decodificatore video numerico riceve i due flussi a 34 Mbit/s e provvede al complesso compito di ricostruire il segnale video numerico, in concordanza con quanto avvenuto nel corrispondente codificatore in trasmissione. Il segnale video decodificato esce ancora in forma numerica, con flusso parallelo a 54 MHz di clock. Dal decodificatore esce anche un flusso a 2 Mbit/s, più precisamente a 2,048 Mbit/s, relativo ai tre segnali audio numerici multiplati.

Si sottolinea che il decodificatore (così come il codificatore) è trasparente: accetta quindi segnali derivanti da entrambi gli standard video HDTV in esperimento.

Il segnale audio numerico a 2 Mbit/s uscente dal decodificatore HDTV Telettra perviene al decodificatore audio MSD-5, dal quale escono tre canali audio analogici in banda base: audio sinistro, audio destro e audio surround.

I singoli segnali audio, attraverso opportuni distributori e amplificatori, sono inviati agli amplificatori di potenza e alle unità di diffusione di sala visione, nonché agli altoparlanti monitoriali della saletta di regia.

Il Bus video numerico proveniente dal decodificatore Telettra è inviato al convertitore D/A. Dopo la conversione si hanno a disposizione i segnali video in banda base in forma di componenti Y, R-Y, B-Y, nonché i sincronismi H e V. Sul convertitore A/D è presente un deviatore tramite il quale l'apparato viene predisposto al trattamento di segnali video a standard 1250/50 oppure 1125/60.

I segnali in componenti vengono immediatamente convertiti in segnali R, G, B per mezzo di due differenti matrici (50 Hz, 60 Hz). Le due matrici si rendono necessarie a causa della diversa colorimetria esistente tra i due standard HDTV da trattare.

I tre segnali R, G, B uscenti dalle matrici costituiscono i segnali video ad alta definizione di fine catena, usufruibili da qualsiasi tipo di display. Come tali vengono distribuiti a tutti gli utilizzatori, per mezzo di equalizzatori/distributori a larga banda (30 MHz).

Dato che i segnali R, G, B sono privi di sincronismi, è necessario provvedere alla sincronizzazione degli utilizzatori per mezzo di segnali di sincronismo separati, nella forma H e V oppure SS. Ciò richiede una ulteriore estensione dell'impianto con l'adozione di due generatori di sincronismi (PG), uno per lo standard 50 Hz ed uno per lo standard 60 Hz, funzionanti in condizione di genlock, ottenuta utilizzando ovviamente i segnali H e V ricevuti. Inoltre all'interno dei PG esiste un generatore di segnali barre e graticcio, nella forma R, G, B, che risultano assai utili per l'allineamento dei proiettori di sala.

Come già detto, la postazione deve poter funzionare indifferentemente a standard 1250/50 e 1125/60: ciò pone l'esigenza di commutare alcune linee video. Il problema è stato risolto con l'adozione di un semplice pannello di commutazione a ponticelli coassiali, commutabili manualmente.

3. Antenna Ricevente

L'antenna (bibl. 2) è formata da un riflettore parabolico con illuminatore centrato a flangia corrugata, che garantisce una elevata discriminazione di polarizzazione.

I parametri principali dell'antenna, che intervengono nel dimensionamento della postazione ricevente, sono il guadagno e la temperatura di rumore. Il guadagno, che esprime il rapporto tra il flusso di potenza dell'antenna in esame e quello del radiatore isotropico, dipende dall'area efficace e dalla frequenza di funzionamento. La temperatura di rumore definisce la densità spettrale del rumore raccolto dall'antenna, al quale contribuiscono il rumore di cielo (rumore cosmico di cielo chiaro e rumore per assorbimento dovuto a eventi atmosferici), che dipende dall'angolo di elevazione dell'antenna ricevente, e il rumore di terra, dovuto all'irraggiamento del suolo.

Il fattore determinante nella scelta dell'antenna resta comunque il diametro del riflettore (che ne determina il guadagno) per il quale non è possibile superare certi limiti per considerazioni di stabilità meccanica e di semplicità di installazione.

Per determinare il diametro dell'antenna ricevente è necessario considerare il fattore di merito G/T della postazione ricevente, rapporto tra il guadagno dell'antenna ricevente e la temperatura di rumore del ricevitore in °K (gradi Kelvin), secondo l'espressione:

$$G/T = G/(T_a + (F - 1) \cdot T_0) \qquad [1]$$

dove G è il guadagno e T_a la temperatura di rumore dell'antenna ricevente, F è la cifra di rumore del ricevitore, $T_0 = 290\,°K$.

Quanto maggiore è il guadagno dell'antenna e quanto minore è il rumore generato nel ricevitore, tanto maggiore è il fattore di merito.

Poiché nella [1] interviene il parametro F, cifra di rumore del ricevitore, che dipende fortemente dalle caratteristiche dell'unità esterna, si passa all'esame della stessa; successivamente, con i parametri ad essa relativi e con il valore di fattore di merito richiesto, sulla base del bilancio del collegamento (bibl. 3) si potrà procedere al calcolo del diametro dell'antenna.

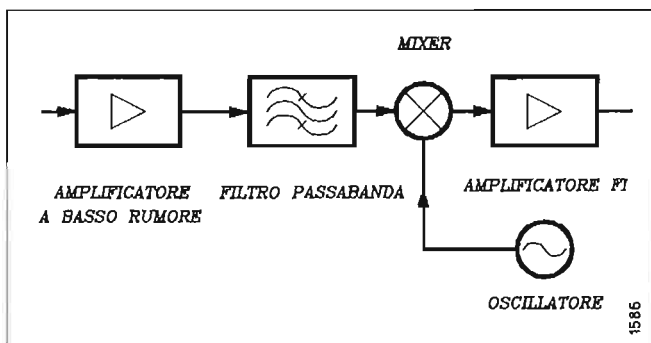



Fig. 2 — Schema a blocchi del convertitore a microonde dell'unità esterna.

4. Unità esterna

L'unità esterna è costituita dal convertitore a microonde, posto immediatamente dietro l'illuminatore dell'antenna, con la funzione di amplificare il segnale ricevuto da satellite e trasferire il medesimo dalla frequenza di ricezione RF a 12 GHz, alla prima frequenza intermedia a 1,4 GHz, dove è introdotta una ulteriore amplificazione.

Il convertitore, rappresentato schematicamente in figura 2, è costituito da un amplificatore a basso rumore, da un mescolatore (o convertitore propriamente detto) con il relativo oscillatore locale, seguito da un amplificatore alla prima frequenza intermedia. Tra l'amplificatore a basso rumore e il mescolatore è inserito un filtro passabanda per l'attenuazione del segnale immagine e per ridurre la reirradiazione dell'oscillatore locale.

I parametri che contraddistinguono il convertitore sono il guadagno, la cifra di rumore, la stabilità della frequenza del segnale convertito al variare delle condizioni ambientali, il rumore di fase.

Il guadagno complessivo è dato dalla somma dell'amplificazione a microonde e di quella a frequenza intermedi, però la prima è di gran lunga la più importante e la più critica. Infatti ogni amplificatore aggiunge al segnale e al rumore presenti al suo ingresso, un rumore interno generato dai dispositivi che lo compongono. Pertanto l'amplificatore a microonde, primo componente del convertitore, deve avere una cifra di rumore bassa e deve amplificare il segnale ricevuto di un valore sufficiente affinché il contributo del rumore generato nella catena a valle sia trascurabile e non degradi la qualità del segnale.

La cifra di rumore dell'amplificatore determina, in massima parte, la cifra di rumore del convertitore completo. Infatti, in una cascata di sottosistemi come quella di figura 2 risulta:

$$[2] \qquad F = F_1 + (F_2 - 1)/G_1 + (F_3 - 1)/(G_1 \cdot G_2)$$

dove F_1 e G_1 sono i parametri dell'amplificatore a microonde; F_2 e G_2 quelli del mixer ($F_2 = L_2$, essendo L_2 la perdita di conversione del mixer), F_3 la cifra di rumore dell'amplificatore FI.

Quindi, tanto più elevato è il guadagno dell'amplificatore a microonde, tanto più la cifra di rumore globale del convertitore si avvicina a quella dell'amplificatore isolato.

Se il guadagno e la cifra di rumore del ricevitore sono parametri tradizionalmente determinanti per il bilancio di un collegamento via satellite, il rumore di fase e la stabilità di frequenza sono parametri specialmente importanti nel caso di segnali numerici a modulazione PSK, in quanto l'informazione è legata alla fase del segnale trasmesso e la variazione di frequenza all'ingresso del demodulatore non deve superare certi limiti prestabiliti.

Perciò, mentre la ricezione da satelliti a diffusione diretta (DDS) di segnali codificati PAL o MAC è ormai possibile su vaste aree, con buona qualità, pur impiegando sistemi riceventi di medie caratteristiche (antenne di diametro 90 cm o meno e unità esterne con 2 dB di cifra di rumore, senza particolari esigenze di stabilità e di rumore di fase) (bibl. 4), nel caso di segnali HDTV a modulazione numerica è necessario avere caratteristiche migliori sia come guadagno e cifra di rumore, ma soprattutto come stabilità della frequenza del segnale e del rumore di fase introdotto nella conversione.

Le unità esterne reperibili sul mercato (bibl. 5) sono in maggior parte prodotte in Giappone, anche se commercializzate in Europa sotto diverse marche, ed hanno caratteristiche adeguate per la ricezione DDS, per la quale è sufficiente una stabilità di frequenza di $\pm 3 \cdot 10^{-4}$ e non vi sono particolari esigenze per il rumore di fase.

Per questo motivo, le unità esterne impiegate nelle po-

Fig. 3 — Antenna ricevente ed unità esterna della postazione ricevente a Palazzo Labia (Venezia).

Tabella 1.
CARATTERISTICHE MISURATE DELLE UNITÀ ESTERNE

Guadagno	:	53 ÷ 55 dB
Cifra di rumore	:	1,3 ÷ 1,4 dB
Rumore di fase SSB a 10 KHz	:	− 81 dBc/Hz max
Offset di frequenza	:	± 1 MHz
Deriva di frequenza		
da − 40 a + 50 °C:		1,5 MHz max
da + 20 a + 50 °C:		0,5 MHz max

stazioni riceventi sono state individuate in seguito ad una accurata indagine di mercato, che ha permesso di scegliere, tra le numerose esistenti, quelle meglio rispondenti alle esigenze sopraddette.

È stata fatta una accurata verifica delle caratteristiche con l'impiego di speciale strumentazione di misura. Per la misura della stabilità di frequenza è stato effettuato il monitoraggio continuo, sottoponendo le unità esterne a cicli termici automatici in camera climatica (bibl. 5), per stabilire in termini quantitativi, per ogni esemplare, la deriva di frequenza nell'intervallo di temperatura prevedibile per le condizioni ambientali di impiego.

I risultati delle misure di circa 20 esemplari presentano bassa dispersione, essendo contenuti nei limiti dei valori riportati in tabella 1; si rileva, inoltre, un offset rispetto al valore nominale della frequenza, di entità compresa entro ± 1 MHz. Perciò ogni unità è corredata delle specifiche misure ad essa relative, per consentire di recuperare tale deviazione mediante i circuiti all'uopo predisposti nell'unità interna.

Considerate le caratteristiche delle stazioni trasmittenti (bibl. 6) e quelle del pay-load su satellite, la valutazione del bilancio del collegamento (bibl. 3) per trasmissione immune da errori richiede un rapporto portante/rumore, relativo alla tratta di discesa, corrispondente ad un fattore di merito della postazione ricevente G/T di 20,7 dB/°K.

Pertanto, con i dati dell'unità esterna suddettti, è ora possibile, dalla [1], ricavare il guadagno dell'antenna ricevente.

Risulta:

[3] $$G = G/T \cdot (Ta + (F - 1) \cdot T_0)$$

da cui, per G/T = 20,7 dB/°K, T_0 = 290 °K, T_a = 85 °K, F = 2 dB (1,4 dB è la cifra di rumore del convertitore, 0,6 dB sono le perdite introdotte dalle connessioni) si ricava G = 44,7 dB, corrispondente, assumendo una efficienza di 0,6, ad una antenna con riflettore di diametro 1,75 metri, per cui si assume come valore definitivo 1,8 metri. è da notare che questo valore deriva dalla necessità di avere un margine molto ampio per garantire l'efficienza del collegamento, con altissimo grado di probabilità, anche in caso di condizioni meteorologiche avverse.

Le misure effettuate (bibl. 7) e soprattutto il perfetto e corretto funzionamento delle postazioni riceventi per tutta la durata dei Campionati mondiali, anche con forte temporale con grandine, confermano la validità di quanto detto.

A riprova sta il fatto che, nel punto di ricezione presso il *Centro Ricerche* RAI, è stata usata sperimentalmente anche un'antenna del diametro di 70 cm con cielo chiaro; si aveva ancora un margine di circa 3 decibel.

5. Unità interna

Durante la fase preparatoria di studio delle postazioni riceventi e di sperimentazione preliminare, è stata esaminata la possibilità di utilizzare unità interne di tipo commerciale, eventualmente apportando le opportune modifiche: questa soluzione avrebbe avuto il vantaggio di consentire un risparmio non indifferente in termini di tempo e di risorse umane.

Però nessuno degli apparati esaminati si è rivelato adatto al particolare tipo di segnale da ricevere ed è stato di conseguenza indispensabile, da parte del *Centro Ricerche* della RAI, affrontare dapprima la progettazione e la costruzione di un prototipo, da sottoporre ad una serie di test che ne dimostrassero la piena funzionalità e rispondenza alle caratteristiche richieste, e in seguito la costruzione e la taratura di una piccola serie di 10 esemplari.

Durante il periodo di progettazione e sperimentazione iniziale in laboratorio, si è rivelato indispensabile l'utilizzo del «simulatore di satellite», realizzato precedentemente dal *Centro Ricerche* della RAI. Il simulatore ha consentito di riprodurre in un banco di misura, con ottima approssimazione, come si è riscontrato in seguito, le condizioni presenti sul collegamento reale via satellite per quanto riguarda il rumore e le distorsioni.

I severi requisiti imposti in sede di progettazione dell'unità interna dalla modulazione QPSK a 70 Mbit/s erano i seguenti:

— elevata larghezza di banda, in modo che su tutto il collegamento, fino all'ingresso del demodulatore, l'unica limitazione fosse in pratica quella costituita dal filtro I-MUX del transponder del satellite Olympus (banda di 42 MHz a − 3 dB);
— ridottissimo ritardo di gruppo in una banda di almeno 35 MHz;
— elevata stabilità di frequenza e minimo rumore di fase dell'oscillatore locale di conversione.

In base a tali requisiti è stata realizzata l'Unità Interna RAI ST/RS tipo 9305, della quale nel seguito vengono descritte le funzioni svolte, le caratteristiche e le prestazioni, che sono riportate in tabella 2.

La figura 4 mostra la vista frontale dell'apparato, mentre il suo circuito a blocchi è rappresentato in figura 5.

5.1 Conversione di frequenza

La funzione principale dell'unità interna è la conver-

Tabella 2.
PRINCIPALI CARATTERISTICHE DELL'UNITÀ INTERNA

Frequenza d'ingresso (1ª FI)	:	1418,62 MHz
Livello d'ingresso minimo	:	− 46 dBm
Livello d'ingresso massimo	:	− 26 dBm
Impedenza d'ingresso	:	50 Ω
Frequenza di uscita (2ª FI)	:	140 MHz
Livello di uscita	:	− 13 dBm
Impedenza di uscita	:	75 Ω
Guadagno	:	33 dB
Risposta di ampiezza nella banda di ± 30 MHz	:	± 0,5 dB max
Ritardo di gruppo nella banda di ± 17,5 MHz	:	5 ns max

Fig. 4 — Vista frontale dell'Unità Interna RAI ST/RS tipo 9305.

sione di frequenza del segnale proveniente dall'unità esterna: il segnale dalla 1ª FI (1418,62 MHz) viene convertito alla 2ª FI (140 MHz).

La conversione è effettuata in un unico passo, con oscillatore locale ad una frequenza 140 MHz più bassa del segnale entrante, in modo che non vi sia inversione dello spettro.

Poiché nella modulazione QPSK l'informazione trasmessa è associata alle quattro diverse fasi che può assumere il segnale modulato, ai fini delle prestazioni del sistema ricevente è essenziale che nel processo di conversione vi sia un contributo trascurabile del rumore di fase da parte dell'oscillatore locale: l'oscillatore utilizzato nell'unità interna in oggetto ha buone caratteristiche a questo riguardo ed è un oscillatore sintetizzato con VCO e PLL, realizzato dalla ditta Teko-Telecom. Il suo rumore di fase su singola banda laterale è migliore di − 85 dBc/Hz a 10 KHz: con questo valore le prestazioni del ricevitore non sono degradate in modo apprezzabile.

Come è stato detto nel capitolo 4, le unità esterne utilizzate per le postazioni riceventi per HDTV numerica sono di tipo commerciale per DDS e, pur essendo già state accuratamente misurate e selezionate, presentano un offset della frequenza di uscita (1ª FI) che varia, rispetto al valore nominale, da esemplare ad esemplare e risulta al massimo di ± 1 MHz (tabella 1); per di più la frequenza dipende anche dalle variazioni termiche, con una variazione massima di 0,5 MHz passando da + 20 °C a + 50 °C (tabella 1).

A sua volta il successivo demodulatore QPSK accetta

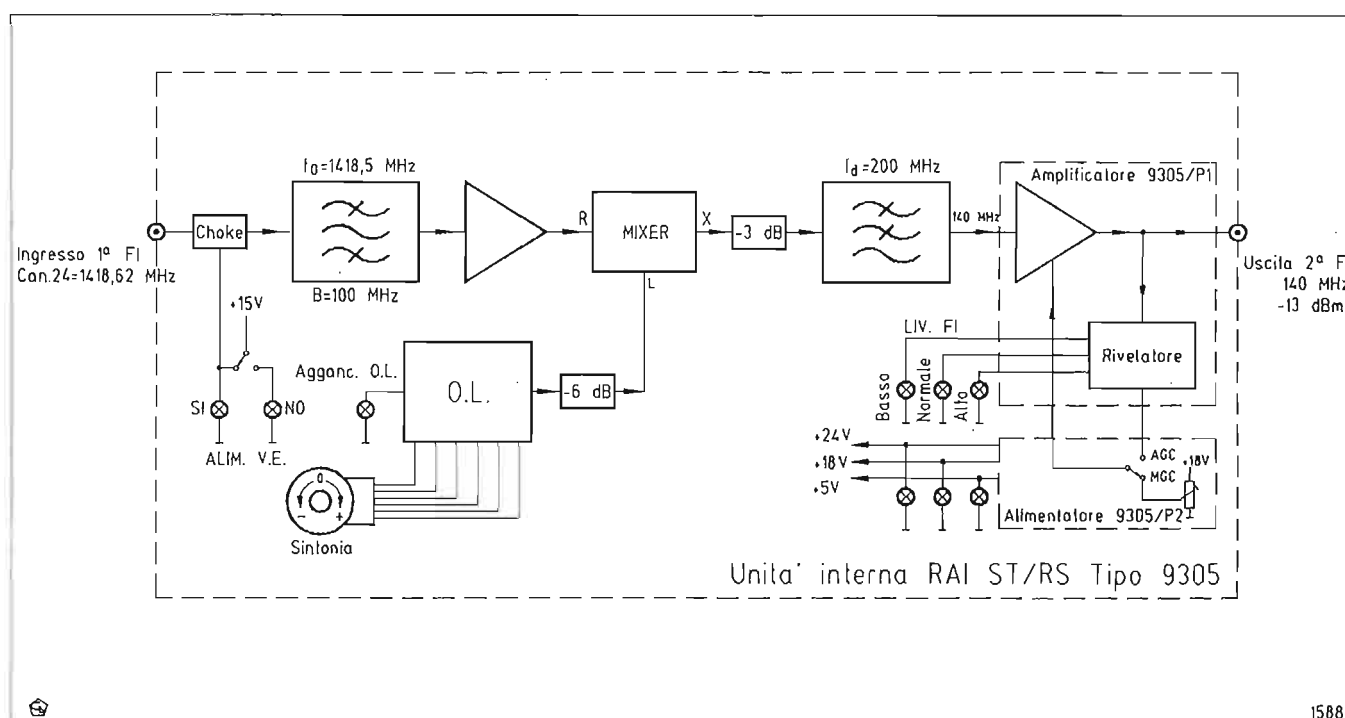



Fig. 5 — Schema a blocchi dell'Unità Interna RAI ST/RS tipo 9305.

variazioni della frequenza del segnale d'ingresso che non devono superare ± 1 MHz, secondo quanto viene indicato dalla casa costruttrice e confermato da prove di laboratorio.

Di conseguenza è risultato necessario fornire l'unità interna di un comando che consenta di compensare opportunamento, nell'ambito indicato, le variazioni di frequenza introdotte dall'unità esterna.

La soluzione adottata, preferita ad altre per ragioni sia di urgenza, sia di praticità di impiego, è stata la predisposizione di un commutatore che, intervenendo manualmente sulla frequenza dell'oscillatore locale, consente di variare la frequenza a scatti di 500 KHz, fino ad un massimo di ± 6 MHz rispetto al valore nominale (140 MHz).

La corretta regolazione viene fatta leggendo su un contatore la frequenza della portate ricostruita all'interno del demodulatore QPSK, il cui valore è lo stesso del segnale modulato d'ingresso: in questo modo la frequenza del segnale all'ingresso del demodulatore può essere contenuta entro uno scarto massimo di 250 KHz dal valore nominale.

È tuttavia da notare che, in pratica, per tutta la durata dell'esperimento non è stato necessario alcun intervento manuale.

5.2 Alimentazione della unità esterna

L'unità interna è in grado di alimentare l'unità esterna attraverso il connettore d'ingresso e il cavo coassiale di discesa.

È stata scelta allo scopo la tensione di ± 15 V, che viene fornita e mantenuta stabilizzata da un regolatore di tensione.

Con tale tensione è anche possibile alimentare amplificatori supplementari, che servivano a compensare la maggiore attenuazione introdotta da ulteriori spezzoni di cavo, nell'eventualità che in qualche installazione i 50 metri di cavo previsti in sede di progetto risultassero insufficienti.

Un deviatore permette di inserire, od escludere, la tensione di alimentazione all'unità esterna: le due condizioni sono segnalate da indicatori a LED.

5.3 Amplificazione

La determinazione del guadagno complessivo dell'unità interna è stata fatta in base ai seguenti parametri:

— minimo livello d'ingresso alla 1ª FI (che dipende dal flusso a terra nel punto di ricezione più sfavorito, dal diametro dell'antenna e dal guadagno dell'unità esterna);
— attenuazione del cavo coassiale tra unità esterna e unità interna;
— livello d'ingresso al demodulatore QPSK.

L'amplificazione richiesta all'unità interna deve anche compensare le perdite di inserzione dei filtri e la perdita di conversione del mixer ed è ripartita tra la 1ª FI, con un amplificatore che copre l'intera banda 0,95 ÷ 1,75 GHz con guadagno di circa 17 dB, e la 2ª FI, con un amplificatore di uscita a 140 MHz con guadagno di circa 30 dB.

In questo modo l'amplificazione tra l'ingresso alla 1ª FI e l'uscita a 140 MHz è di circa 33 dB, che assicurano alcuni decibel di margine rispetto al valore normalmente necessario.

Il livello di uscita a 140 MHz dell'unità interna deve essere mantenuto costantemente al valore di − 13 dBm su 75 Ohm, richiesto all'ingresso del successivo demodulatore QPSK, compensando sia le variazioni del segnale ricevuto dal satellite, sia, per le differenze fra le caratteristiche dei vari punti riceventi, i diversi guadagni delle unità esterne e le differenti attenuazioni dei cavi coassiali.

A questo scopo, l'amplificatore a 140 MHz comprende un circuito di controllo automatico di livello: esso è costituito da un attenuatore variabile a diodi PIN, comandato da un circuito che rivela il livello del segnale di uscita e ne ricava una tensione di errore.

5.4 Allarmi

L'unità interna è fornita di alcuni indicatori a LED che segnalano condizioni di funzionamento anomalo; essi sono:

— assenza di alimentazione all'unità esterna:
— livello RF scarso o eccessivo: le due spie si accendono rispettivamente allorché, in conseguenza di valori non corretti del livello d'ingresso alla 1ª IF, il livello di uscita a 140 MHz scende al di sotto di − 15 dBm o aumenta oltre − 12 dBm. Esiste comunque ancora un margine di sicurezza al di là di questi limiti, in quanto il campo di funzionamento del demodulatore QPSK va da − 10 a − 23 dBm;
— aggancio del PLL dell'oscillatore locale: la spia si accende quando il PLL non è correttamente agganciato, il che comporta instabilità della frequenza di uscita.

5.5 Caratteristiche di ampiezza e di ritardo di gruppo

In sede di progettazione si è tenuto conto che la larghezza di banda dell'intero collegamento deve essere limitata essenzialmente dal filtro I-MUX del transponder del satellite Olympus: circa 42 MHz a − 3 dB.

La larghezza di banda dell'unità interna deve essere maggiore per quanto possibile di questo valore, in modo da non introdurre distorsioni apprezzabili sul segnale numerico nel canale RF, anche in conseguenza degli scarti della frequenza del segnale entrante alla 1ª FI dovuti all'unità esterna, tenendo però presente la necessità di attenuare sufficientemente sia il rumore presente alla frequenza immagine, sia gli eventuali canali, trasmessi da Olympus o da altri satelliti, che possono essere causa di interferenze.

A questo scopo, in base al tipo di modulazione (QPSK) e alla velocità di trasmissione (70 Mbit/s), un compromesso soddisfacente è risultato quello di adottare:

— una larghezza di banda di circa 100 MHz a − 3 dB per il filtro passabanda d'ingresso alla 1ª FI;
— una banda fino a circa 200 MHz entro 0,2 dB per il filtro passa basso alla 2ª FI.

Fig. 6 — Apparati di conversione di frequenza, demodulazione e decodifica audio e video ubicati nella Sede RAI a Palazzo Labia (Venezia).

Con queste caratteristiche dei filtri, il ritardo di gruppo complessivo è risultato contenuto entro 5 ns in 35 MHz.

5.6 PRESTAZIONI

Le prestazioni dell'unità interna si sono rivelate corrispondenti alle aspettative.

Infatti si è riscontrato che il sistema ricevente composto dall'unità esterna e da quella interna, collegate tra di loro con 50 metri di cavo tipo RG214/U, forniva con il simulatore di satellite a parità di condizioni di misura, gli stessi valori di tasso di errore in funzione del C/N misurati con il convertitore dalla 1ª FI a 140 MHz dello stesso simulatore di satellite, che utilizza un generatore RF professionale quale oscillatore locale.

Alle stesse conclusioni si è giunti provando una condizione di maggiore distanza tra unità esterna e unità interna, con 4 spezzoni di cavo 50 emtri ciascuno, collegati l'un l'altro tramite amplificatori da 17 dB (complessivamente 3): anche in questo caso i risultati ottenuti non si discostavano da quelli relativi alla sola coppia modemodulatore più simulatore di satellite.

6. Demodulatore QPSK

La scelta del tipo di mo-demodulatore QPSK da utilizzare (il modulatore per le due stazioni trasmittenti, fissa e mobile, il demodulatore per le postazioni riceventi) ha comportato un notevole lavoro preparatorio, consistente dapprima nel cercare di reperire, tra i modelli di produzione industriale, quelli adatti a funzionare al bit-rate necessario (60 ÷ 70 Mbit/s), poi nell'effettuare misure con il simulatore di satellite per individuare quel modello che assicurava le prestazioni migliori.

Al termine, la scelta è caduta sul mo-demodulatore prodotto dalla ditta tedesca ANT.

Le caratteristiche più significative del demodulatore sono le seguenti:

— Frequenza FI d'ingresso	: 140 MHz
— Tolleranza offset frequenza	: ± 1 MHz
— Livello d'ingresso FI	: − 13 dBm
— Impedenza d'ingresso FI	: 75 Ω
— Filtro FI	: largo
— Tipo carrier recovery	: Costa's loop, collegato con clock recovery e ricerca con rampa
— Tipo clock recovery	: PLL con quarzo e ottimizzazione dell'istante di campionamento
— Filtro di banda base	: radice coseno rialzato, α = 0,5, con equalizzatore di ritardo di gruppo
— Bit-rate di modulazione	: 69,87 Mbit/s (al demultiplex)
— N° uscite digitali	: 2 (dal demultiplex)
— Bit-rate uscite	: 34,368 Mbit/s
— Impedenza di uscita	: 75 Ω
— Tipo di segnale	: HDB3
— Decodifica	: differenziale (nel demultiplex)
— Descrambler	: uno per ogni flusso a 34 Mbit/s (nel demultiplex), di tipo autosincronizzante
— Fattore di moltiplicazione degli errori	: 5 (decodifica differenziale e descrambler)

Per le corrispondenti caratteristiche del modulatore si rimanda all'articolo sulle stazioni trasmittenti contenuto in questa stessa pubblicazione (bibl. 6).

Le prestazioni del mo-demodulatore, in loop FI, sono tipicamente un $E_b/N_0 = 9{,}5$ dB necessario per un tasso di errore di 10^{-4}, contro un valore ideale di 8,4 dB.

7. Decodificatore video

Il decodificatore come il codificatore RAI/Telettra è oggetto di un articolo su questa stessa pubblicazione (bibl. 8).

8. Convertitore Digitale/Analogico

I convertitori A/D e D/A saranno oggetto di un articolo su questa pubblicazione.

9. Decodificatore audio

Si richiedeva la trasmissione di un segnale audio stereofonico e di un segnale audio monofonico (quest'ultimo per l'effetto «surround»), tutti di alta qualità.

Poiché sul co-decodificatore HDTV era disponibile un flusso a 2 Mbit/s per la trasmissione dei canali musicali, sono stati adottati gli apparati del sistema tipo MSD-5 della Telettra.

Gli apparati sono multiplex e demultiplex numerici adatti alla trasmissione di un massimo di 5 canali audio musicali monofonici, con sistema PCM, ad un bit-rate di 2,048 Mbit/s: il codice di linea è di tipo HDB3, in accordo con la Racc. G.703 del CCITT.

I canali musicali hanno una banda di 15 KHz, adottano l'enfasi della Racc. J.17 del CCITT e sono utilizzabili a coppie nel caso di trasmissione di programmi stereofonici.

La velocità di campionamento per ciascun canale è di 32.000 campioni al secondo, con codifica lineare a 14 bit per campione e compressione istantanea a 11 bit, secondo la legge di compressione/espansione a 11 segmenti (Doc. CMTT/186-S): con l'aggiunta del bit di parità, la parola codificata passa a 12 bit e pertanto la velocità di trasmissione di un singolo canale musicale è di 384 kbit/s.

Il demultiplex è provvisto di un sistema di correzione degli errori basato sul bit di parità, con sostituzione dell'ultimo campione corretto al posto di quello errato, fino a 3 campioni errati consecutivi, oltre i quali interviene un particolare sistema di silenziamento graduale che impedisce la generazione di «click» in altoparlante.

10. Matrici

Come già annunciato nel capitolo 2, la presenza delle matrici si rende necessaria in quanto il segnale video HDTV in uscita dal convertitore D/A è nella forma Y, B-Y, R-Y.

Inoltre, poiché le condizioni di matricizzazione nei due standard in esperimento sono diverse, occorrono due apparati differenti.

Le matrici in questione, realizzate dal *Centro Ricerche* RAI constano di alcuni stadi di amplificazione e adattamento di impedenza del segnale video di ingresso, a cui fanno seguito le reti resistive di ricombinazione ed un circuito di uscita amplificatore e traslatore di impedenza, in grado di alimentare con 1 V_{pp} una linea video a 75 Ohm.

Occorre porre molta cura nel realizzare gli stadi amplificatori antecedenti le reti combinatorie. Infatti piccole diversità nei ritardi di propagazione (alcuni nanosecondi) determinano delle esaltazioni o attenuazioni della banda passante alle alte frequenze: ciò in quanto le frequenze video al limite della banda sono dell'ordine dei 30 MHz, a cui corrispondono 33 ns di periodo. Uno sfasamento di soli 3,3 ns tra i segnali dei tre canali determina, dopo i processi di somma o differenza, errori di ampiezza già apprezzabili sui segnali R, G, B.

Altrettanto delicata risulta la scelta delle resistenze di matrice. Queste devono essere di precisione e di valore fisso, onde garantire la stabilità termica e la stabilità nel lungo termine.

In accordo con la Raccomandazione n. 240M dello SMPTE, le espressioni di riferimento per il calcolo delle matrici a 60 Hz sono:

$$Y = (0{,}701 \cdot G) + (0{,}087 \cdot B) + (0{,}212 \cdot R)$$
$$B\text{-}Y = (-0{,}701 \cdot G) + (0{,}913 \cdot B) + (-0{,}212 \cdot R)$$
$$R\text{-}Y = (-0{,}701 \cdot G) + (0{,}087 \cdot B) + (-0{,}788 \cdot R)$$

da queste si ricavano i segnali di crominanza trasmessi:

$$P_b = (-0{,}384 \cdot G) + (0{,}500 \cdot B) + (-0{,}116 \cdot R)$$
$$P_r = (-0{,}445 \cdot G) + (-0{,}050 \cdot B) + (0{,}500 \cdot R)$$

ponendo:

$$P_b = \frac{B - Y}{1{,}826} \qquad P_r = \frac{R - Y}{1{,}576}$$

In accordo con la Raccomandazione n. 601[1] del CCIR, le espressioni di riferimento per il calcolo delle matrici a 50 Hz sono:

$$Y = (0{,}587 \cdot G) + (0{,}114 \cdot B) + (0{,}229 \cdot R)$$
$$B - Y = (-0{,}587 \cdot G) + (0{,}886 \cdot B) + (-0{,}299 \cdot R)$$
$$R - Y = (-0{,}587 \cdot G) + (-0{,}114 \cdot B) + (0{,}701 \cdot R)$$

si ottengono analogamente a prima:

$$C_b = (-0{,}331 \cdot G) + (0{,}500 \cdot B) + (0{,}169 \cdot R)$$
$$C_r = (-0{,}419 \cdot G) + (-0{,}081 \cdot B) + (0{,}500 \cdot R)$$

ponendo:

$$C_b = \frac{B - Y}{1{,}77} \qquad C_r = \frac{R - Y}{1{,}40}$$

1: All'atto dell'esperimento, il riferimento della colorimetria europea (50 Hz) era ricavato dalla Raccomandazione CCIR 601.

Dalle suddette equazioni si ricavano pertanto le espressioni dei segnali R, G, B da inviare al display.

11. Generatori di sincronismi HDTV 50 Hz e 60 Hz

Fin dall'inizio delle sperimentazioni sui sistemi HDTV, i generatori di sincronismo (PG) sono stati un elemento notevolmente critico, in modo particolare per lo standard 50 Hz. Infatti, non esistendo normative certe, tali apparati non erano disponibili sul mercato.

Presso il Centro Ricerche della RAI si è pertanto provveduto da tempo allo studio e sviluppo di generatori di sincronismi HDTV a 50 Hz e a 60 Hz, con caratteristiche di versatilità tali da poter sopperire alle molteplici necessità di uso e da permettere adattamenti alle eventuali variazioni dello standard specifico.

Le caratteristiche più interessanti di questi PG sono:

- generazione dei segnali di base: H, V, Blk V, Blk H;
- generazione dei segnali di SS, sia nella forma tradizionale di Bisinc (polarità solo negativa), sia nella forma specifica di Trisinc (doppia polarità);
- possibilità di genlock con segnale di riferimento Trisinc o H e V separati;
- possibilità, in condizione di genlock, di spostamento della fase dei segnali generati rispetto alla fase del segnale di riferimento;
- possibilità di doppio andamento delle fasi dei segnali H, V, SS, onde ottenere una tessitura di immagine a 1250 Hz interallacciato oppure 625 Hz progressivo;
- possibilità di usufruire in RGB di un segnale di barre e di un segnale di graticcio generati internamente.

Il segnale di sincronismo denominato Trisinc è un nuovo segnale ideato espressamente per la televisione ad alta definizione. Esso sostituisce il tradizionale SS (e suoi derivati quali il Black Burst) della televisione tradizionale. Trattasi quindi di un segnale di sincronismo composito, di tipo analogico, portatore contemporaneamente della informazione di riga H e di semiquadro V con livello negativo di 0,3 V. Pertanto è perfettamente adatto alla formazione di segnali video compositi di 1 V_{pp}. Il suo andamento completo è a ciclo bipolare (vedi figura 7). Tale caratteristica costituisce la novità essenziale: infatti da essa scaturiscono le sue peculiarità funzionali.

I principali vantaggi offerti dal Trisinc sono:

1) Possibilità di utilizzare il Trisinc come un normale SS tradizionale. Questo in quanto esiste la possibilità di allineamento automatica del picco negativo, con la conseguente eliminazione delle eventuali distorsioni di inviluppo e la successiva semplice separazione del sincronismo dal segnale video.
2) Superiore garanzia di giusta individuazione del sincronismo, utilizzando il primo fronte negativo come preset ed il successivo fronte positivo come sincronismo.
3) Possibilità di utilizzare il Trisinc come un normale SS individuando semplicemente il primo fronte o fronte negativo. Questa particolarità permette di sincronizzare di H anche con il segnale Trisinc alcuni apparati progettati per sincronismi tradizionali.
4) Esatta individuazione della fase del sincronismo, utilizzando come riferimento l'istante di transito per lo zero del secondo fronte, o fronte positivo.

Fig. 7 — Rappresentazione grafica e caratteristiche del segnale di sincronismo Trisinc.

Quest'ultima possibilità risulta essere la più interessante in quanto da essa derivano i maggiori vantaggi, cioè:

- minore sensibilità al rumore grazie alla possibilità di filtraggio del segnale (riduzione del jitter);
- scarsa sensibilità alle distorsioni lineari del canale di trasmissione, in quanto l'istante esatto di sincronismo è individuato dal passaggio per lo zero di un segnale simmetrico;
- indifferenza alle distorsioni non lineari, in quanto inesistenti nell'istante di zero.

Da quanto illustrato deriva la importante possibilità di poter utilizzare con perfetta fasatura reciproca un segnale video presentato nella forma di tre segnali compositi in componenti o di tre segnali compositi R, G, B.

12. Distributori/Equalizzatori

L'esigenza di distribuire, all'interno dell'impianto ricevente, il medesimo segnale video su più linee e contemporaneamente di compensare le perdite in banda introdotte dai cavi di cablaggio e collegamento (vedi figura 1), ha richiesto l'utilizzatore di opportuni distributori/equalizzatori in grado di trattare segnali HDTV. Questi apparati sono stati prodotti dalla ditta italiana Elpro, in collaborazione con gli specialisti del Centro Ricerche RAI. Tale collaborazione ha permesso

Fig. 8 — Sale visione presso l'ex Stabilimento della FIAT «Lingotto» a Torino.

di ottenere una ottimizzazione del prodotto, sia dal punto di vista elettrico che meccanico, e con essa una elevata versatilità e affidabilità. Affidabilità confermata dalla totale assenza di avarie durante tutto il periodo dell'esperimento.

Le caratteristiche funzionali di tali apparati sono così riassumibili:

— possibilità di compensare linee di lunghezza variabile da 0 a 100 metri di cavo coassiale RAI (il cavo coassiale RAI e sostanzialmente un cavo di tipo RG59, ma con doppia calza schermante);
— banda passante estesa da 25 Hz a 30 MHz entro $\pm 0,2$ dB con 3 carichi da 75 Ohm, corrispondenti al 50% del carico massimo;
— ampiezza dei segnali in transito di 0,7 V_{pp}, se segnali di sola luminanza, oppure segnali di 1 V_{pp} se completi di sincronismi. Tali ampiezze, inoltre, sono accettabili per segnali fino al limite della banda e in qualsiasi condizione di equalizzazione;
— impedenza di uscita bassa, e costante al variare della frequenza, fino a contenere le variazioni del livello del segnale di uscita entro $\pm 0,5$ dB per qualsiasi frequenza e per variazioni di carico da una a sei linee a 75 Ohm.
— possibilità di trattare anche segnali di sincronismo al massimo livello di standard (4 V_{pp}).

13. Conclusioni

Le postazioni riceventi allestite presso varie città italiane per dimostrare, in prima mondiale, il nuovo sistema di HDTV numerica a 70 Mbit/s per collegamenti punto-multipunto, hanno richiesto l'impegno congiunto di un gruppo di specialisti del *Centro Ricerche* RAI e della fattiva collaborazione di importanti industrie nazionali per progettare e realizzare gli apparati speciali per il trattamento del segnale numerico.

Le stazioni riceventi, sebbene costituite in gran parte da apparati sperimentali o da prototipi realizzati in tempi molto stretti, hanno operato in modo assolutamente soddisfacente ed affidabile, senza alcun disservizio in oltre 40 giorni di attività, anche in condizioni meteorologiche avverse.

(3927)

BIBLIOGRAFIA

1. - M. Ardito, G.F. Barbieri, M. Cominetti: *ITALIA '90: una prima mondiale di collegamento HDTV numerico via satellite OLYMPUS*, in questo stesso numero.
2. - G.P. Pacini: *Antenne riceventi per diffusione diretta da satellite*, «Elettronica e Telecomunicazioni», n. 2 e 3, 1989.
3. - M. Cominetti, A. Morello: *Trasmissione numerica punto-multipunto via satellite di segnali HDTV*, in questo stesso numero.
4. - G.P. Pacini: *Satellite Olympus-F1*, «Elettronica e Telecomunicazioni», n. 2 e 3, 1989.
5. - M. Ariaudo, F. Cuccia, P. Forni: *Convertitori a microonde per la ricezione da satellite: disponibilità commerciale, caratteristiche misurate, stabilità in frequenza*, «RAI - Centro Ricerche», Relazione Tecnica n. 89/15/I.
6. - G. Moro, B. Sacco, D. Tabone: *ITALIA '90: Stazioni trasmittenti di segnali HDTV numerici via satellite*, in questo stesso numero.
7. - G. Garazzino, A. Morello, B. Sacco: *Prove di trasmissione di segnali HDTV a 70 Mbit/s via OLYMPUS dalla stazione del CPTV di Roma*, «RAI - Centro Ricerche», Relazione Tecnica n. 90/01/I.
8. - M. Barbero, S. Cucchi: *ITALIA '90: Codifica del segnale televisivo numerico*, in questo stesso numero.

ITALIA '90: SISTEMA DI PROIEZIONE HDTV

D. TOGNETTI, S. DEL CONT BERNARD*

SOMMARIO — L'articolo illustra le tecniche realizzative del proiettore HDTV Seleco, utilizzato per visualizzare le riprese televisive in Alta Definizione delle partite del Campionato Mondiale di Calcio '90. L'articolo intende evidenziare inoltre i fattori tecnologici sui quali si interverrà nell'immediato futuro per incrementare ulteriormente le prestazioni del sistema.

SUMMARY — ***HDTV Projection System.*** *- The paper illustrates the implementation techniques of the Seleco HDTV projector, used to display some shootings of the football matches during the 1990 Worldcup in Italy. The article aims at pointing out the technological factors to be improved in the near future in order to further increase the sistem performance.*

1. Introduzione

Nel settore della videoproiezione, dominato finora da grandi aziende multinazionali (Sony, Panasonic, General Electric), sta rapidamente emergendo SELECO, la società elettronica italiana che, grazie alla sua attività nell'ambito dei progetti europei (in particolare il progetto Eureka per l'Alta Definizione) e ad una ormai consolidata collaborazione con i laboratori del Centro Ricerche RAI, intende proporsi come azienda di riferimento in Europa in questo settore dei videoproiettori che, nel caso dell'Alta Definizione, rappresenteranno verosimilmente il primo prodotto ad essere immesso sul mercato.

2. La proiezione HDTV su grande schermo

Durante lo svolgimento del campionato mondiale di calcio 1990, le riprese effettuate in Alta Definizione, sono state visualizzate, presso le sale visione RAI predisposte a Milano, Venezia, Perugia e Napoli, da proiettori televisivi ad Alta Definizione della SELECO S.p.A.

I sistemi di proiezione rappresentano attualmente una modalità fondamentale per la presentazione di immagini televisive su schermi di grande formato (oltre i 100 pollici di diagonale). È noto infatti che le dimensioni dei tubi a raggi catodici (CRT) sono limitate da fattori tra i quali principalmente la resistenza meccanica alla pressione esercitata dall'atmosfera e il conseguente peso del vetro. Pertanto immagini di grande formato possono essere ottenute con un CRT soltanto mediante un ingrandimento ottico e quindi con un sistema di proiezione. Poiché le immagini in Alta Definizione sono apprezzate maggiormente su grandi schermi, ne consegue la naturale vocazione dei sistemi di proiezione per la visualizzazione delle informazioni in HDTV.

Il sistema di proiezione, da un punto di vista elettronico, è sostanzialmente un triplo monitor: tre immagini, nei colori fondamentali rosso, verde e blu vengono prodotte su tre CRT di piccolo formato ed alta luminosità. Le tre immagini, presenti sul piano del fosforo dei CRT, vengono focalizzate da un sistema di lenti sul piano dello schermo secondo le consuete leggi dell'ottica geometrica; sullo schermo, che possiede proprietà riflettenti, le immagini si sovrappongono esattamente producendo l'immagine a colori percepita dallo spettatore.

Si comprende quindi che i blocchi fondamentali che concorrono all'ottenimento della luminosità e della risoluzione richieste per l'Alta Definizione sono quattro:

1) l'elettronica di pilotaggio del CRT;
2) il CRT;
3) l'ottica di proiezione;
4) lo schermo riflettente.

In particolare la luminosità della immagine sullo schermo L_d (in c_d/m^2) è così esprimibile:

$$L_d = G\, L_t \frac{T}{4\, F^2(1 + M^2)}$$

in cui:

- G è il guadagno dello schermo rispetto allo schermo lambertiano;
- L_t è la luminosità dell'immagine sul CRT in c_d/m^2;
- T è il coefficiente di trasmissione dell'ottica di proiezione;
- F è f/A (rapporto tra lunghezza focale e apertura);
- M è l'ingrandimento del sistema di lenti.

La risoluzione viene invece caratterizzata mediante una

* Ing. Dante Tognetti Direttore Ricerca e Sviluppo della SELECO S.p.A.; Ing. Stefano Del Cont Bernard Responsabile Settore Videoproiezione della SELECO S.p.A.
Dattiloscritto pervenuto alla Redazione il 12 settembre 1990.

funzione di trasferimento MTF (modulation transfer function) che esprime la percentuale di rapporto di contrasto nell'immagine di uscita (sullo schermo) rispetto all'immagine di ingresso (sul CRT). La funzione viene quantificata con percentuale in corrispondenza a date frequenze spaziali costituite da linee alternativamente bianche e nere (rapporto di contrasto 100%). L'MTF complessivo di un proiettore TV è il prodotto dei singoli MTF lungo la catena ottica: CRT, lenti e schermo. Il risultato finale, inteso come massima frequenza del segnale visualizzabile sullo schermo, sarà dato dalla cascata della funzione di trasferimento elettrica degli amplificatori video e dell'MTF della catena ottica.

Per il CRT è consuetudine esprimere la risoluzione in termini di larghezza della traccia riferita ai punti in cui essa possiede una luminosità pari al 50% del massimo (ricordiamo che il profilo di luminosità di una traccia, in direzione perpendicolare a quella della traccia stessa, è approssimativamente di tipo gaussiano). Come valore indicativo si osserva che una traccia con larghezza di 80 micrometri può essere considerata equivalente ad un MTF del 60% alla frequenza spaziale di 5 l_p/mm (5 linee pari per mm significa 5 linee nere e 5 bianche alternate su un intervallo di 1 mm), oppure una risoluzione limite di 12 l_p/mm (frequenze superiori non sono visibili).
L'MTF di un sistema di lenti per proiezione ad alta risoluzione si aggira su valori del 50% a 10 l_p/mm, mentre per gli schermi i valori di MTF sono normalmente superiori sia ai corrispondenti valori per i CRT che per le lenti. Da queste considerazioni e dall'analisi di quanto tecnologicamente disponibile in tema di componenti, emerge che l'elemento fondamentale e punto debole di tutta la catena è rappresentato dal CRT, mentre l'elettronica di pilotaggio e l'ottica, raggiungono le prestazioni richieste per ottenere una immagine in alta definizione. Ulteriore fatto di debolezza del CRT è la forte dipendenza della sua MTF dalla luminosità dell'immagine, con il risultato di perdita di definizione in presenza di immagini ad elevato contrasto e luminosità, condizione spesso richiesta con i proiettori TV, per compensare l'effetto della luce ambientale diffusa.

Il compito della SELECO S.p.A. è stato quindi quello di sviluppare la parte elettronica in accordo con le specifiche richieste per i sistemi in HDTV, mentre per quanto riguarda CRT, lenti e schermi sono stati avviati contatti con le aziende leader in campo mondiale, nei singoli settori, al fine di individuare i componenti con le prestazioni richieste ed eventualmente definire le specifiche per componenti di tipo custom.

3. Parte elettronica

Su questo versante si è svolto la gran parte del lavoro progettuale della SELECO. I blocchi fondamentali possono essere così elencati:

— canale video (amplificatori) per ciascuna componente RGB;
— processore di sincronismi;
— stadi di potenza per le deflessioni orizzontale e verticale;
— sistema di correzione geometrica N-S ed E-W;
— sistema di convergenza dinamica (peculiare dei proiettori televisivi);
— sistema di focalizzazione elettrica dinamica.

a questi blocchi si deve aggiungere il sistema separato di generazione dell'alta tensione (EHT), gli alimentatori e il sistema distribuito di protezione contro i guasti (con particolare attenzione all'alta tensione).

I tre canali video RGB sono a loro volta costituiti da tre parti: un buffer differenziale d'ingresso avente il compito di eliminare eventuali componenti di modo comune, un amplificatore intermedio su cui agiscono il controllo di contrasto e di luminosità, e un amplificatore finale con massima tensione picco-picco di uscita di 110 V. L'ingresso è di tipo RGB con ampiezza 0,7 volt.

Lo standard HDTV europeo indica in 27 MHz la larghezza di banda del segnale (con periodo orizzontale di scansione di 32 microsecondi). Poiché il proiettore SELECO ha come modalità operativa anche la scansione orizzontale a 16 microsecondi (per impiego con segnali HDTV a frequenza di quadro di 100 Hz), la banda video (a − 3 dB) necessaria risulta essere di 54 MHz. Dal momento che il segnale elettrico viene tradotto in segnale ottico sul piano del fosforo ed il segnale ottico subisce, prima di arrivare allo spettatore, una ulteriore degradazione dovuta alla MTF del sistema ottivo di proiezione, si è convenuto di adottare amplificatori con banda maggiore di 54 MHz. Attualmente la larghezza di banda impiegata nella catena video, dall'ingresso RGB al catodo del CRT, è di 80 MHz a − 3 dB.

Il processore di sincronismi svolge il compito di pilotare gli stadi di potenza delle deflessioni, avendo come ingresso i segnali di sincronismo. Al fine di permettere la massima flessibilità di impiego è stata prevista la possibilità di applicare all'ingresso i segnali di sincronismo nei seguenti formati:

1) separato H-V, con livello TTL e polarità arbitraria;
2) composito TTL con polarità arbitraria;
3) composito analogico;
4) composito analogico sul canale del verde.

La selezione tra i vari formati viene svolta automaticamente. Il processore di sincronismi inoltre prevede la possibilità di funzionamento con frequenza orizzontale di scansione compresa tra 14 kHz e 70 kHz e frequenza verticale tra 40 Hz e 110 Hz in modo da permettere l'impiego non solo per il segnale HDTV ma anche per l'interfacciamento con le schede grafiche di computer e workstation.

Gli stadi di potenza per le deflessioni ricalcano schemi circuitali ormai ampiamente collaudati. La deflessione orizzontale è di tipo risonante serie con transistor interruttori di tipo MOS. Su ciascun giogo orizzontale la corrente picco-picco per la deflessione a pieno raster (angolo max di 70°) risulta essere di circa 8 ampere con un impulso di flyback di 600 volt. La deflessione verticale è realizzata con amplificatori di potenza integrati di tipo audio.

La correzione di geometria N-S ed E-W è realizzata mediante forme d'onda lineari e paraboliche con controllo separato sui quattro semiquadri: alto, basso, destra e sinistra.

La convergenza elettronica è fondamentale nei sistemi di proiezione ed ha lo scopo di portare in perfetta sovrapposizione le immagini proiettate dai CRT sullo schermo, infatti in assenza di convergenza elettronica le tre immagini, che sono proiettate da tre punti diversi, non potrebbero sovrapporsi sullo schermo e pertanto si rende necessaria una predistorsione geometrica delle immagini dei CRT laterali (rosso e blu) in modo che appaiano corrette, dopo la proiezione ottica sghemba dei CRT laterali, rispetto al CRT verticale (verde).

Il tipo di realizzazione adottata va sotto il nome di convergenza analogica ed è caratterizzata dalla possibilità di effettuare, con notevole facilità, le correzioni di convergenza, agendo indipendentemente sui quattro quadranti delle immagini R e B, con predistorsioni di tipo lineare e parabolico.

La focalizzazione nel proiettore HDTV SELECO è di tipo elettrostatico e richiede tensioni di circa 10 kV. In realtà il pennello elettronico, con tensione di focalizzazione costante, non risulterebbe uniformemente focalizzato su tutto lo schermo, in quanto l'ottica elettronica del CRT produce una superficie di focalizzazione in prima approssimazione sferica e quindi non coincidente con il piano del fosforo. Per ottimizzare la focalizzazione si è adottata la tecnica della modulazione della tensione di fuoco mediante forme d'onda quadratiche e ampiezze di modulazione di 400 volt ottenendo in tal modo una focalizzazione uniforme su tutto lo schermo.

La tensione di accelerazione del CRT viene generata mediante un apposito alimentatore in alta tensione. Il suo funzionamento è indipendente dalla scansione di riga in quanto quest'ultima non è fissa, ma può avere frequenze comprese in un campo di più di due ottave, in dipendenza del tipo di segnale video. La tenzione di accelerazione è di 32 kV.

Particolare attenzione è stata dedicata infine alle protezioni contro i guasti, al fine di garantire l'immunità dell'elettronica in corrispondenza a condizioni irregolari di funzionamento. Il sistema distribuito di protezione attua pure una semplice funzione di autodiagnosi per la rapida individuazione di guasti accidentali.

4. Il cinescopio

Il CRT impiegato possiede una diagonale di 9" e raster utile di 7"; la dimensione 9" risulta essere la maggiore praticamente disponibile in quanto vincolata dall'apertura ottica del sistema di lenti per proiezione. Dimensioni maggiori del CRT richiederebbero ottiche custom molto costose e con dimensioni meccaniche notevoli.

Poiché l'immagine HDTV ha un formato 16:9 ne consegue che le dimensioni massime dell'immagine possono essere 155 · 87 mm. La visualizzazione di un segnale a 27 MHz su un tempo di riga attiva di 26,67 microsecondi equivale quindi a circa 4,7 l_p/mm. Pertanto il solo cinescopio avente un MTF del 70% a 4,7 l_p/mm rappresenta un passabasso con frequenza di taglio a 27 MHz.

Purtroppo questi valori di risoluzione risultano essere, allo stato attuale, difficilmente compatibili con le esigenze di luminosità richieste per l'immagine proiettata. Si tratta di un problema peculiare dei cinescopi per proiezione a causa delle piccole dimensioni che implicano contemporaneamente aumento di luminosità ed aumento di risoluzione dello spot in ragione inversa alle dimensioni della diagonale. È opportuno sottolineare che in un monitor per HDTV a visione diretta, lo spot possiede dimensioni notevolmente maggiori e luminosità di molto inferiore al corrispondente spot di un cinescopio per proiezione, si comprende quindi come sia tecnicamente difficile proporre un paragone diretto tra una immagine su cinescopio da proiezione e su cinescopio da visione diretta.

Il CRT impiegato da SELECO possiede una larghezza di traccia al 50% valutabile in 120 micrometri. La tensione anodica impiegata è di 32 kV e la focalizzazione è di tipo elettrostatico. Il CRT possiede una propria cella di raffreddamento a convezione libera riempita con un liquido otticamente trasparente.

Fig. 1 — Sala visione allestita al Centro Congressi della Camera di Commercio a Perugia. Si notino, affiancati, i due proiettori per Alta Definizione (uno dei quali di scorta) ed, in primo piano, il proiettore PAL di emergenza.

5. L'ottica di proiezione

L'ottica di proiezione è di tipo rifrattivo, cioè si tratta di un insieme di lenti con correzione cromatica. Tra la prima lente del gruppo ottico e la superficie del CRT vi è un liquido organico, con opportuno indice di rifrazione, avente lo scopo di compiere un «adattamento» tra i mezzi di propagazione della radiazione ottica. In altre parole il liquido riduce le perdite per riflessione che, per un sistema di proiezione, sono molto dannose in quanto conducono alla generazione di luminosità diffusa sulla superficie del cinescopio abbassando il massimo rapporto di contrasto ottenibile. Il coefficiente di trasmissione, con accoppiamento a liquido, risulta essere dell'80%.

L'MTF dell'obiettivo risulta essere del 50% a 10 l_p/mm su tutto il campo.

Altri due parametri interessanti dell'ottica impiegata sono la lunghezza focale di 167 mm (con ingrandimento di 25, l'ingrandimento varia da 14 a 45) e il parametro f/A avente il valore 1,15.

Allo stato attuale l'ottica rappresenta un elemento critico non tanto per i limiti tecnologici quanto per i costi di produzione degli obiettivi ad alta risoluzione. I costi maggiori sono legati agli obiettivi che utilizzano lenti asferiche in vetro; recentemente costi più contenuti sono stati ottenuti mediante l'impiego di lenti realizzate con materie plastiche.

Tuttavia gli obiettivi da proiezione per alta risoluzione trovano ancora un mercato molto ristretto e in più casi la disponibilità è legata a costosi progetti di tipo custom.

6. Lo schermo di proiezione

Lo schermo rappresenta un elemento chiave per il sistema di visualizzazione, la sua funzione è di raccogliere l'immagine proiettata dall'ottica e rifletterla verso lo spettatore. È evidente che la luminosità dell'immagine dipende in modo diretto dal coefficiente di riflessione della superficie dello schermo, tuttavia il comportamento dello schermo non è dettato solo dalle sue capacità riflettenti, ma soprattutto dalle proprietà diffondenti.

Uno schermo ottimale dovrebbe permettere la visione, con la stessa luminosità, sotto un qualsiasi angolo orizzontale compreso tra + 90° e − 90° rispetto all'asse dello schermo, mentre l'angolo verticale dovrà coprire tutte le possibili posizioni degli spettatori (tipicamente un angolo di 30°).

Quindi la diffusione deve essere direttiva e ciò comporta la possibilità di elevare la potenza ottica diffusa verso lo spettatore, riducendo la corrispondente potenza nelle direzioni non utilizzate. Vi sono schermi aventi guadagni anche sensibilmente superiori all'unità e che producono immagini molto luminose nella direzione dell'asse dello schermo a scapito della luminosità in direzioni molto periferiche, guadagni attualmente raggiungibili si aggirano sul valore 10.

Purtroppo schermi molto direttivi riproducono male la purezza del colore, l'immagine infatti è realizzata dalla sovrapposizione di tre immagini nei colori fondamentali e tali immagini sono proiettate da tre direzioni diverse, la direttività quindi attribuisce a ciascun colore un diverso guadagno dipendente anche dalla posizione dello spettatore e alterando il colore in modo diverso nei diversi punti dello schermo. Il peso del difetto diminuisce comunque con l'angolo di visione, pertanto gli schermi molto direttivi sono egualmente usati dal momento che per essi l'angolo di visione è ovviamente stretto (tipicamente 60°).

La scelta del tipo di schermo è quindi dettata da più fattori e tra questi la distribuzione del pubblico è forse la principale. Infatti dalla suddetta distribuzione discende sia la dimensione dello schermo necessario, sia i massimi angoli di visione e quindi la direttività ed il guadagno dello schermo.

Per l'impiego nelle proiezioni di *Italia '90* si è deciso di impiegare uno schermo a basso guadagno e largo angolo di visione (oltre 120°).

La dimensione della superficie di proiezione è di 140 pollici con un rapporto di 16:9 mentre la distanza tra proiettore e schermo è di circa 4,5 metri.

Fig. 2 — Sala visione a Perugia

7. Conclusioni

Il sistema della videoproiezione rappresenta attualmente una valida tecnica per la visualizzazione di immagini in Alta Definizione grazie alla sua peculiarità di riprodurre immagini su schermi giganti, peculiarità che per prima permetterà l'apertura del mercato HDTV.

Sono inoltre stati evidenziati margini per incrementare ulteriormente le prestazioni, non tanto sul versante circuitale che è ormai sostanzialmente consolidato, quanto sul versante dei cinescopi da proiezione, delle ottiche e degli schermi. In particolare i cinescopi rappresentano il collo di bottiglia: ad essi sarà rivolta particolare attenzione nella futura attività di ricerca assieme allo sviluppo di gruppi ottici a basso costo.

A seguito di questa prima fase di sperimentazione sono stati raggiunti risultati molto incoraggianti e vi sono tutte le premesse per sviluppi, nel prossimo futuro, di grande interesse, sia dal punto di vista tecnologico che di mercato. (3909)

ITALIA '90: COLLEGAMENTI IN FIBRA OTTICA PER HDTV

G. B. GREBORIO - V. SARDELLA*

SOMMARIO — Il problema di trasferire segnali HDTV su una distanza di alcuni chilometri è stato risolto con collegamenti in fibra ottica monomodale. È stato utilizzato un cavo che fa parte della rete dei collegamenti di contributo della RAI ed è stato posato un cavo aggiuntivo di lunghezza circa 300 m per completare il collegamento. L'articolo descrive le soluzioni adottate per il trasferimento di un segnale numerico formato da due flussi a 34 Mbit/s e di un segnale HDTV analogico le cui componenti sono state trasmesse in tecnica SCM (SubCarrier Multiplexing). Le problematiche e le prestazioni di quest'ultimo collegamento vengono analizzate più in dettaglio.

SUMMARY — ***Optical Fiber Links for HDTV*** *- The problem of transferring HDTV signals over a distance of some kilometres was solved by monomodal optical fiber links. These links were implemented using an optical fiber cable, pertaining to the RAI contribution network, and an additional cable (approximately 300 m long) laid down on purpose. This article describes the solutions adopted for transferring a digital signal consisting of two 34 Mbit/s streams, and an analogue HDTV signal, whose components were transmitted using the SCM (SubCarrier Multiplexing) technique. All the problems and the performance of this last link are analysed more in detail.*

1. Introduzione

Il *Punto di immissione* situato nei pressi dello Stadio Olimpico di Roma è stato collegato al Centro di produzione TV, punto di trasmissione verso il satellite Olympus, per mezzo di fibre ottiche monomodali.

Per questa via sono stati inoltrati il segnale numerico in direzione Punto di immissione - Centro di produzione TV per la trasmissione verso il satellite Olympus ed il segnale video analogico, ricevuto da satellite, in direzione inversa, Centro di produzione TV - Punto di immissione, per controllo.

La figura 1 riporta lo schema a blocchi dei collegamenti.

Il segnale numerico uscente dal codificatore HDTV e formato da due flussi a 34 Mbit/s è stato inviato su due fibre ottiche modulando direttamente in intensità la portante ottica.

Il segnale video monitorio, disponibile in banda base sotto forma di componenti (Y, P_R, P_B) e sincronismi (H, V), è stato trasmesso su altre due fibre utilizzando la tecnica SCM (SubCarrier Multiplexing) con due sistemi a quattro canali ciascuno e convertitori elettro-ottici a banda larga operanti nella seconda finestra ottica. Il sistema era in grado anche di trasmettere i segnali audio utilizzando inseritori - disinseritori audio a 7,5 MHz 4PSK collegati a uno dei canali ancora disponibili.

La scelta di queste soluzioni, non ottimali dal punto di vista dello sfruttamento della capacità di trasmissione dei collegamenti in fibra ottica, è stata dettata non da particolari esigenze tecniche legate alla trasmissione di segnali HDTV, ma dalla struttura dei segnali da trasferire, dalla disponibilità immediata di apparati numerici a 34 Mbit/s e analogici a 4 canali ciascuno e dall'avere a disposizione una certa quantità di fibre ottiche libere.

2. Il collegamento ottico

Lo stadio Olimpico di Roma ed il Centro di produzione TV della RAI sono collegati da un cavo ottico monomodale che fa parte della rete dei collegamenti di contributo della RAI. Questi speciali collegamenti, destinati esclusivamente ad usi interni, sono generalmente equipaggiati con apparati progettati dal *Centro Ricerche* RAI adatti a trasferire segnali audio e video di qualità, ad esempio da studi esterni al punto di messa in onda (bibl. 1 e 2).

Il collegamento Punto di immissione — Centro di produzione TV per l'invio dei segnali HDTV è stato realizzato utilizzando alcune fibre ottiche disponibili nel cavo Stadio Olimpico — CP/TV prolungandole dalla terminazione lato Stadio Olimpico fino al Punto di immissione con cavi speciali posati appositamente.

La lunghezza di questo collegamento aggiuntivo era di circa 300 m ed è stato necessario utilizzare cavi ottici con particolari caratteristiche meccaniche (elevata resistenza a tiro, percussione e schiacciamento) perché il percorso di posa comprendeva alcuni tratti in tubazione, altri aerei, altri su pavimento o terreno con raggi di curvatura ridotti (a volte inferiori al valore indicato dal costruttore del cavo) e protezioni limitate.

(*) Dott. Giovanni Battista Greborio e ing. Vincenzo Sardella del Centro Ricerche RAI (Torino) - Laboratorio Altafrequenza.
Dattiloscritto pervenuto alla Redazione il 21 agosto 1990.

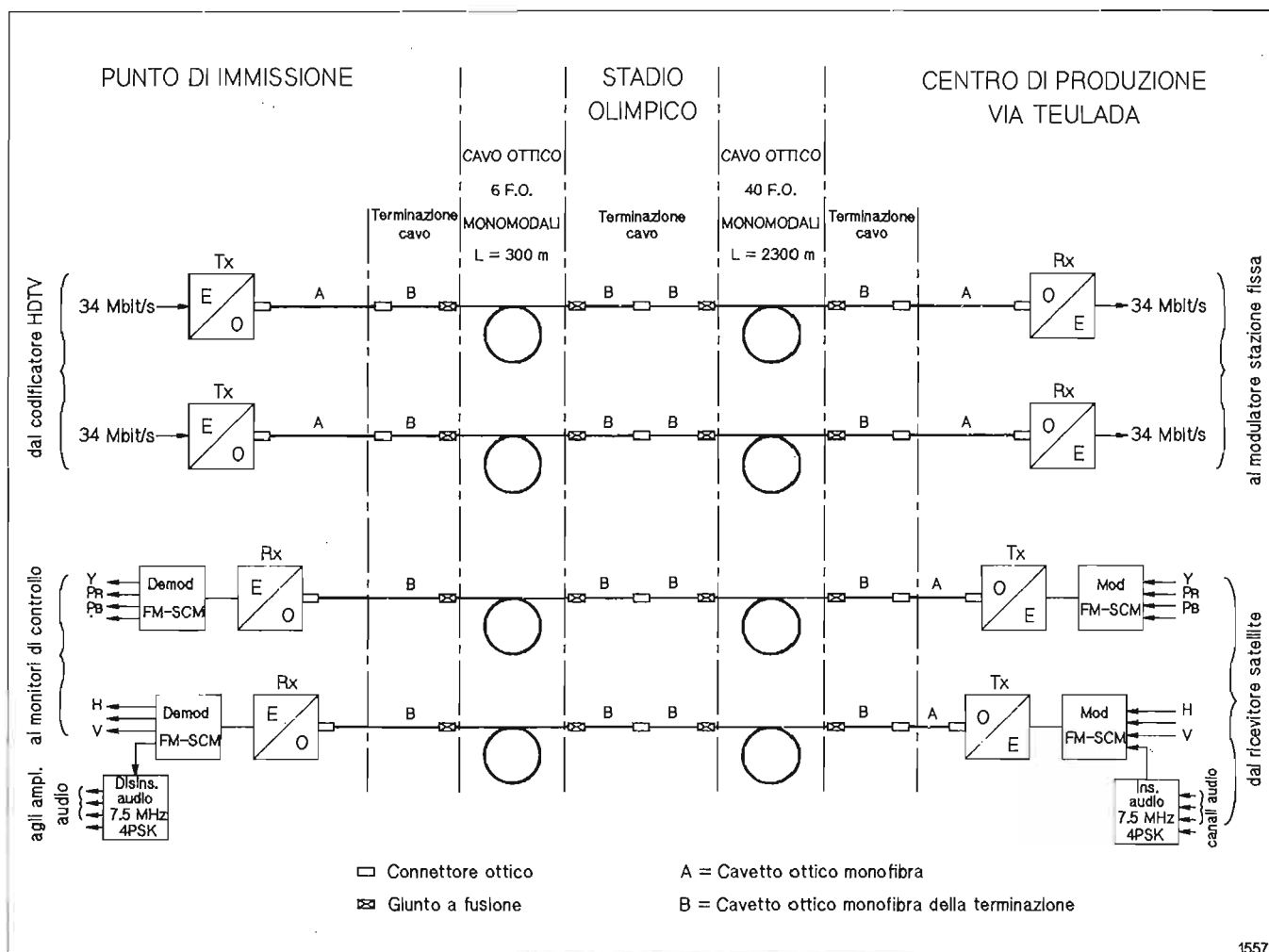



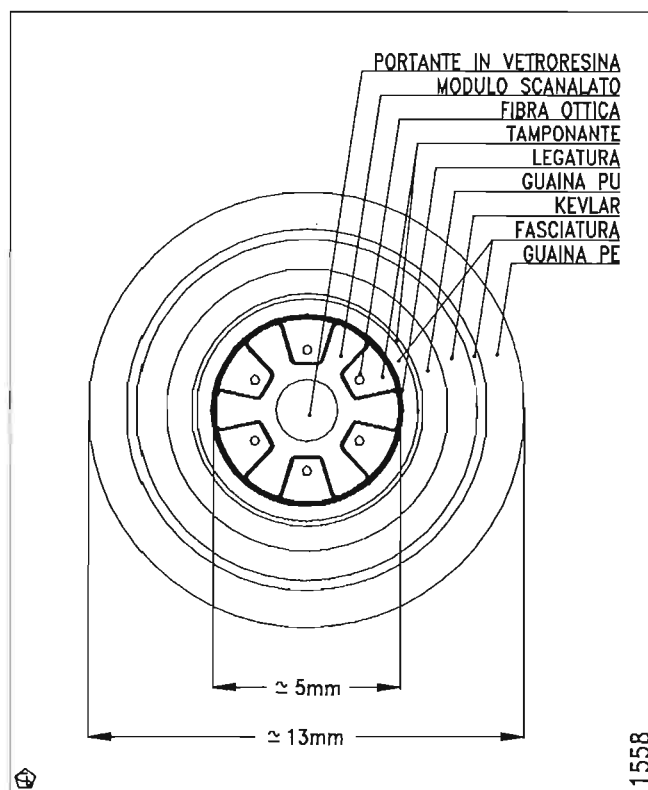
Fig. 1 — Schema a blocchi dei collegamenti in fibra ottica per HDTV.

La scelta è caduta sul cavo ottico a nucleo scanalato tipo PKE a 6 fibre ottiche monomodali interamente dielettrico e autoportante prodotto dalla CEAT CAVI, il cui schema costruttivo è mostrato in figura 2 e le cui principali caratteristiche sono riportate in tabella 1.

Sono stati posati due cavi, uno principale ed uno di riserva, ma nonostante la particolarità dell'impianto e le conseguenti sollecitazioni dovute alla posa, non si sono

Tabella 1.
PRINCIPALI CARATTERISTICHE DEL CAVO OTTICO

N. fibre	6
Tipo fibre	monomodali (Racc. G652 CCITT)
Attenuazione fibre	0,35 dB/km a $\lambda = 1{,}31$ μm
Diametro esterno	13,7 mm
Raggio min. di curvatura	25 cm
Forza max. di trazione	250 kg
Schiacciamento	≤ 2400 kg (IEC 794-1-E3)
Percussione	≤ 12 kg × m (IEC 794-1-E3)

Fig. 2 — Schema costruttivo del cavo ottico a 6 fibre tipo PKE della CEAT CAVI.

avuti danneggiamenti o avarie e quindi non si è dovuto fare ricorso al cavo di riserva.

Prima della installazione degli apparati terminali sono state effettuate una serie di misure ottiche per verificare le attenuazioni delle fibre utilizzate e la bontà dei giunti e dei connettori presenti sul collegamento nei riguardi sia della attenuazione che della riflessione. A titolo di esempio la figura 3 riporta la curva di retrodiffusione di un collegamento CP/TV - Punto di immissione e ritorno.

3. Collegamento HDTV numerico Punto di immissione - CP/TV

Per trasmettere il segnale numerico HDTV sono state utilizzate due fibre ottiche monomodali equipaggiate ciascuna con il sistema di trasmissione TL 73 della Siemens Telecomunicazioni (bibl. 3).

Il trasmettitore ottico è del tipo a laser in 2ª finestra con una potenza ottica media in uscita di − 13 dBm. Esso svolge le funzioni di convertitore di codice (da HDB3 a NRZ) e generatore dei bit di parità per il controllo della qualità di trasmissione, effettua lo scrambling ed infine la conversione elettro-ottica del segnale.

Il ricevitore ottico svolge le funzioni complementari al trasmettitore rivelando il segnale ottico tramite un fotodiodo a valanga (APD). La sua sensibilità è di − 42,5 dBm (per un tasso d'errore di 10^{-9}) ed il suo campo di AGC è di circa 20 dB di potenza ottica. Dal momento che l'attenuazione totale del collegamento ottico era di circa 2 dB è stato necessario inserire all'ingresso del ricevitore un attenuatore ottico da 10 dB.

4. Collegamento HDTV analogico monitorio CP/TV - Punto di immissione

Per la trasmissione dei segnali monitori disponibili sotto forma di componenti analogiche (Y, P_R, P_B), dei sincronismi (H, V) e dei canali audio in sottoportante si è fatto uso della tecnica FM-SCM, che consiste nel modulare in frequenza più portanti RF con i segnali di banda base ed accorpare i segnali così ottenuti per formare un multiplex di frequenza che successivamente modula in intensità la sorgente ottica; in ricezione vengono effettuate le operazioni inverse (bibl. 4 e 5).

Gli apparati utilizzati sono stati modem a quattro canali allocabili nella banda 40 ÷ 580 MHz della ditta Teko Telecom e convertitori elettro-ottici della ditta Pirelli Focom.

Dovendo trasmettere sei segnali è stato necessario ricorrere a due sistemi identici collegati su due fibre ottiche.

I modem sono stati derivati da un sistema multicanale, prodotto per il normale video composito, e modificati per estendere la banda base a 30 MHz. Le caratteristiche principali di ciascun canale sono:

Banda passante	10 Hz ÷ 30 MHz entro ± 1 dB
Deviazione di frequenza	± 15 MHz/1,5 V_{pp}
Frequenza di uscita	40 ÷ 580 MHz
Larghezza di banda IF	120 MHz
S/N video non pesato	≥ 50 dB in banda 30 MHz (P_{ric} = − 25 dBm)

I convertitori elettro-ottici, opportunamente modificati per adattarli alla particolare applicazione, operano alla lunghezza d'onda di 1,31 μm ed hanno le seguenti caratteristiche:

Banda passante	20 ÷ 600 MHz entro ± 1 dB
Potenza ottica Tx	≈ 0,8 mW in F.O. SM 9/125 μm
Potenza ottica Rx	5 ÷ 300 μW

Per collegamenti analogici con segnali multiplati in frequenza, le non linearità del sistema generano prodotti di intermodulazione che possono cadere entro le bande di frequenza utilizzate. Per questo motivo le portanti RF nei due collegamenti sono state scelte tenendo conto delle non linearità delle sorgenti ottiche per minimizzare il numero dei prodotti di intermodulazione che cadono all'interno dei vari canali e che danno origine a segnali di battimento in banda base.

L'allocazione delle portanti è risultata la seguente:

Collegamento 1	Collegamento 2
f_1 = 155 MHz	f_1 = 160 MHz
f_2 = 310 MHz	f_2 = 319 MHz
f_3 = 425 MHz	f_3 = 425 MHz
f_4 = 541 MHz	f_4 = 530 MHz

Fig. 3 — Curva di retrodiffusione di un collegamento ottico CP/TV - Punto di immissione - CP/TV:

A: fibra ottica CP/TV - Stadio Olimpico;
B: connessione Stadio Olimpico (1 connettore + 2 giunti a fusione);
C: fibra ottica Stadio Olimpico - Punto di immissione;
D: connessione al Punto di immissione con un'altra fibra ottica (1 connettore + 2 giunti a fusione).

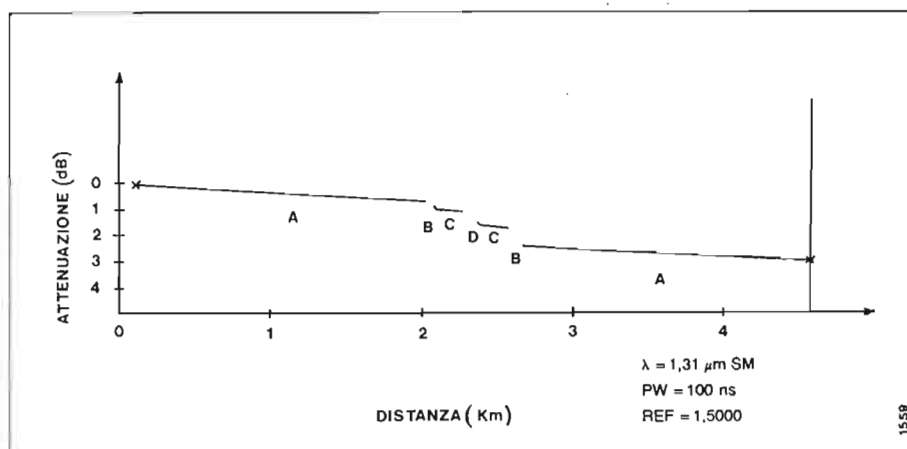
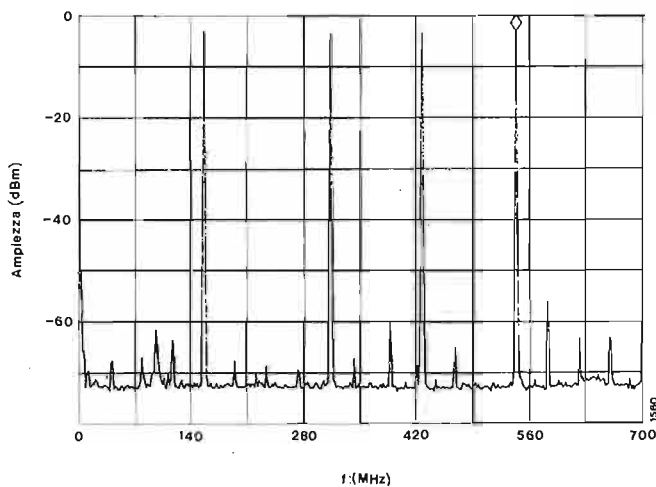



Fig. 4 — Spettro del segnale uscente dal modulatore.

Fig. 6 — Spettro del segnale entrante nel demodulatore relativo alla portante del canale 2.

In laboratorio sono stati caratterizzati tutti i canali disponibili, anche se poi, nella specifica applicazione, non tutti i canali sono stati utilizzati.

Le figure 4, 5, 6 e 7 mostrano gli spettri dei segnali nei vari punti del collegamento 1 simulato su banco nel caso di portanti non modulate.

Confrontando le figure 4 e 5, si può vedere che i prodotti di intermodulazione presenti all'ingresso del demodulatore sono in numero e di ampiezza molto maggiore di quelli presenti all'uscita del modulatore.

Le figure 6 e 7 relative alla portante numero 2 ($f_2 = 310$ MHz) mostrano come, grazie all'opportuna scelta delle frequenze portanti, i prodotti di intermodulazione cadano solo nell'intorno della portante e siano in numero molto limitato.

Nelle figure 8, 9, 10 e 11 sono mostrati gli spettri di banda base dei 4 canali con tutte le portanti non modulate; in esse si possono notare alcuni segnali di battimento sovrapposti al rumore di fondo.

Per ciascun collegamento, è poi stato ancora ottimizzato il livello del segnale elettrico che va a modulare la sorgente ottica e che influisce sia sul rapporto S/N video ottenibile in ciascun canale (tramite la profondità di modulazione della potenza ottica) sia sull'ampiezza dei prodotti di intermodulazione.

Nel collegamento realizzato su banco, con gli apparati utilizzati poi per il collegamento 1, sono stati misurati i seguenti rapporti segnale/rumore non pesato, nella banda 0 ÷ 30 MHz, in presenza di modulazione su tutti i canali:

Canale	S/N (dB)
1	40
2	41
3	40
4	44

Questi valori sono determinati sia dal rumore vero e proprio che dai prodotti di intermodulazione che, dopo demodulazione, cadono entro il canale di bassa frequenza.

Fig. 5 — Spettro del segnale entrante nel demodulatore.

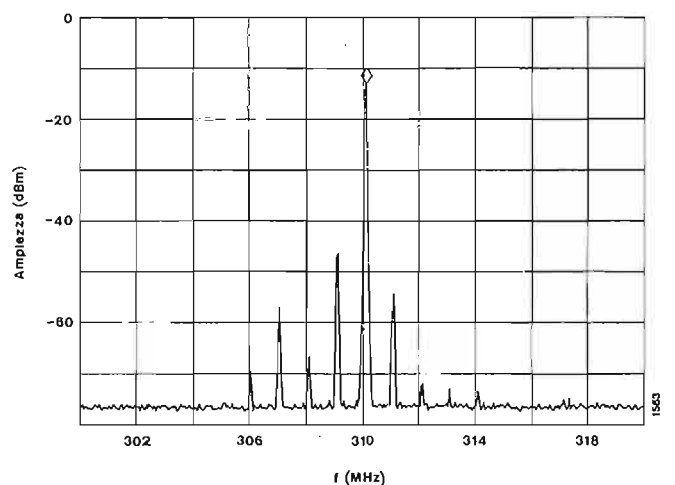



Fig. 7 — Particolare dello spettro di Fig. 6 nell'intorno della portante del canale 2.

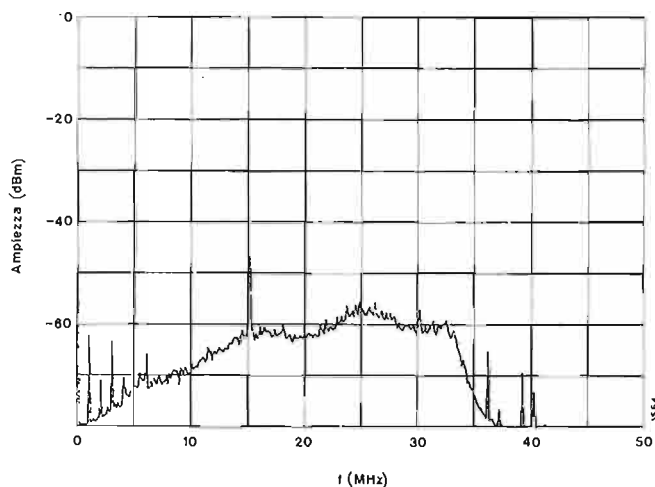



Fig. 8 — Spettro di rumore in banda base del canale 1 senza segnale video applicato.

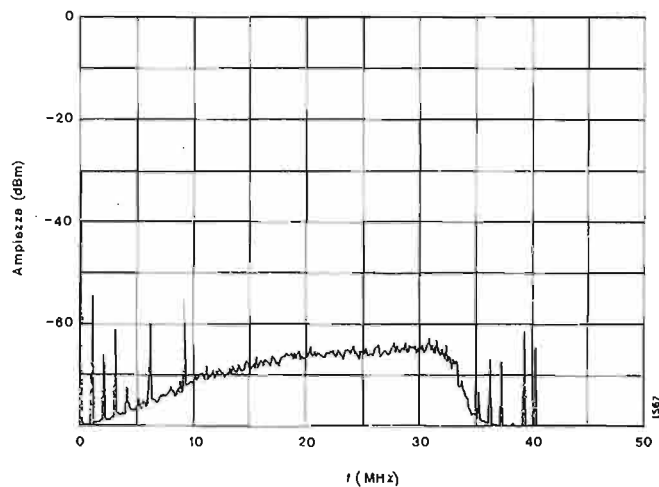



Fig. 11 — Spettro di rumore in banda base del canale 4 senza segnale video applicato.

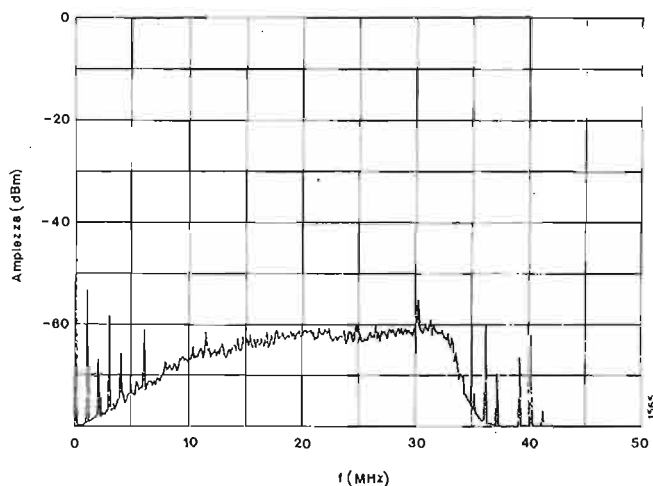



Fig. 9 — Spettro di rumore in banda base del canale 2 senza segnale video applicato.

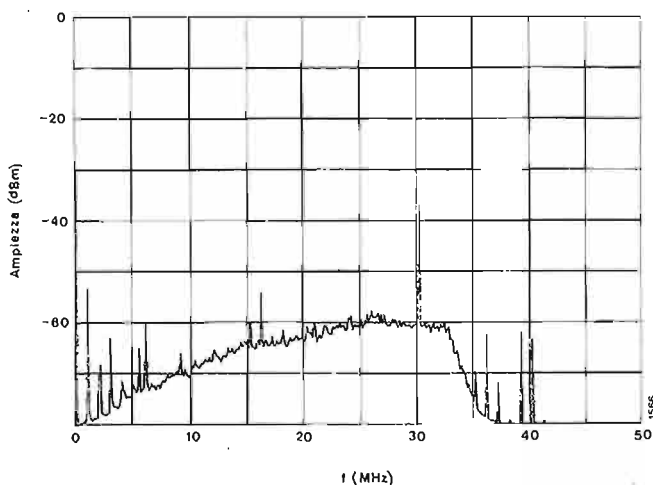



Fig. 10 — Spettro di rumore in banda base del canale 3 senza segnale video applicato.

5. Conclusioni

Le soluzioni adottate per il trasferimento su fibra ottica dei segnali HDTV tra il «punto di immissione», situato presso lo stadio Olimpico, e il Centro di Produzione TV della RAI di Roma si sono rilevate adeguate alle esigenze della sperimentazione dal punto di vista della qualità dei collegamenti e della affidabilità e sicurezza dell'impianto.

La scelta delle tecniche di trasmissione è stata dettata dal formato del segnale HDTV numerico, dalla presenza presso il CP/TV di una catena di ricezione da satellite con uscite analogiche, dal tipo di apparati immediatamente disponibili e dal non avere limitazioni sul numero di fibre ottiche da utilizzare.

Ringraziamenti

Gli autori desiderano ringraziare quanti, colleghi e non, hanno collaborato e si sono adoperati per realizzare in tempi brevi questi collegamenti.

BIBLIOGRAFIA

1. - V. Sardella: *Collegamenti televisivi di contributo su fibre ottiche.* «Elettronica e Telecomunicazioni», n. 2, 1987, p. 85.

2. - G.B. Greborio, V. Sardella: *Collegamenti in fibra ottica S.M. a 70 MHz - apparati tipo 9270.*« Manuale d'istruzione RAI», Aprile 1990.

3. - AA.VV.: *Sistema di trasmissione su fibra ottica a 35 Mbit/s TL73.* «Manuale tecnico Siemens Telecomunicazioni», 11/89.

4. - E. Afasar: *Analog Multichannel TV Transmission on Optical Fibres.* «16[th] International TV Symposium - CATV Session» - Montreux, 17-22 June 89.

5. - F.V.C. Mendis, P.A. Rosher: *CNR Requirements for subcarrier-multiplexed multichannel video FM transmission in optical fibre.* «Electronics letters», 1989, vol. 25, p. 72.

CODIFICA DEL SEGNALE TELEVISIVO NUMERICO: TECNICHE DI COMPENSAZIONE DEL MOVIMENTO ASSOCIATE ALLA DCT IBRIDA

MARIO MURATORI*

SOMMARIO — La tecnologia attuale permette di realizzare codificatori per la riduzione della ridondanza del segnale televisivo, sia a definizione normale che ad alta definizione, particolarmente efficienti in termini di compressione della quantità di informazione del segnale codificato ed in termini di qualità dell'immagine codificata grazie anche all'applicazione di tecniche di compensazione del movimento. Nell'ambito del progetto europeo EU 256, presso il Centro Ricerche RAI si è svolto un lavoro di analisi delle tecniche di compensazione del movimento proposte in sede di normativa internazionale e si è sviluppato un nuovo metodo di codifica dell'informazione associata al movimento che si è dimostrato particolarmente efficiente in termini di codifica e robusto contro gli errori di canale, risultando pertanto adatto alle applicazioni di riduzione della ridondanza studiate nel corso del progetto stesso. In questo lavoro vengono illustrati i risultati salienti di tale analisi, e viene descritto il metodo di codifica sviluppato presso il Centro Ricerche RAI sia in termini di struttura generale del metodo, sia in una sua particolare realizzazione, dimostratasi la più efficiente tra le molte prese in considerazione.

SUMMARY — ***Coding of digital TV signal: techniques of motion compensation, associated with hybrid DCT.*** *Highly efficient codecs for the bit-rate reduction of the television and the HDTV signals are now-a-days feasible because of the advanced present electronic technology, which allows motion compensation techniques to be implemented, thus obtaining high compression ratios and high quality of the coded images. The RAI Research Centre has been involved in the European Project EU 256 in the last years, one of the tasks being to analyze the motion compensation techniques proposed in the international standardizing bodies. Since those techniques didn't result satisfactory enough, a new method for encoding the motion information has been developped which showed to be very efficient and robust against the channel errors, and which therefore is considered to be suitable to be adopted in bit-rate reduction applications. In this article the results of the analysis performed on the proposed motion compensation techniques are shown and the new encoding method is described, both in general form and in one of the different possible implementations which showed to be the most efficient.*

1. Introduzione

Caratteristica tipica delle immagini televisive è la loro variabilità temporale, la quale spesso è associata al movimento reale dei soggetti o degli oggetti ripresi dalla telecamera o al movimento di questa rispetto alla scena inquadrata.

In questi e altri casi la rappresentazine televisiva della scena o di parti di essa appare, da un quadro all'altro, traslata sullo schermo nel corso della sequenza televisiva e appare quindi in «movimento» sullo schermo.

Qui di seguito, con «movimento» si intende più particolarmente il fenomeno per cui parti di immagine si trovano in un certo istante traslate rispetto alla posizione occupata, sullo schermo, in quadri precedenti o seguenti, componenti una medesima sequenza.

In tale definizione non rientrano perciò i fenomeni per cui tali parti di immagine si trovano ruotate o «zoomate», cioè ingrandite o rimpicciolite, o presentano un deformazioni geometriche o di forma in quadri successivi.

Il movimento di tali parti di immagine ha la caratteristica di essere bidimensionale, infatti si realizza sullo schermo televisivo, e viene quantificato e rappresentato con un vettore bidimensionale che, generalmente, indica in ampiezza e in direzione dove la parte di immagine in oggetto si trova, rispetto alla posizione attuale, in un quadro precedente; tale vettore viene chiamato «vettore movimento» ed il suo significato è illustrato in figura 1.

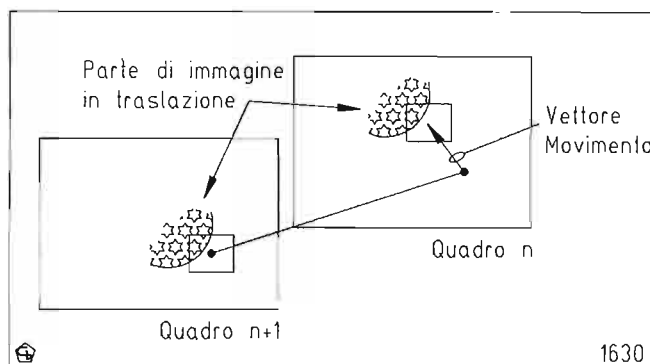



Fig. 1 — Illustrazione del significato di «vettore movimento».

* Ing. Mario Muratori del Centro Ricerche RAI - Torino.
Dattiloscritto pervenuto alla Redazione il 10 gennaio 1991.

In associazione con i vettori movimento, risulta necessario definire l'intervallo di tempo intercorrente tra l'immagine attuale e quella in cui viene ricercata la parte di immagine che ha subito lo spostamento che si vuole misurare; in genere vengono considerati semiquadri o quadri televisivi adiacenti, corrispondenti a intervalli di tempo di 20 ÷ 40 ms rispettivamente nello standard europeo a 50 Hz, ma per speciali applicazioni possono essere considerati anche intervalli di maggior durata e/o i vettori movimento relativi ad imagini successive anziché precedenti.

Le parti di immagine di cui si desidera misurare il movimento sono definite in modo diverso a seconda dell'applicazione. In campo televisivo normalmente si preferisce definirle come blocchi di forma rettangolare di dimensioni che possono, a titolo di esempio, spaziare dai 4 · 4 elementi (4 pixel per 4 linee) ai 32 · 32 elementi.

Diversissime sono le tecniche impiegate per la stima del movimento chiamata in inglese «motion estimation» e di seguito indicata con «ME» (bibl. 5 ÷ 13).

Particolarmente adatte alla suaccennata suddivisione in blocchi sono le tecniche chiamate di «block matching» che misurano quanto un blocco traslato sia simile (da cui l'uso della parola «matching») al blocco considerato appartenente all'immagine di riferimento.

In letteratura sono descritti molti metodi cosiddetti «veloci» per eseguire la stima del movimento, sviluppati per diminuire il carico e il tempo di calcolo rispetto a quelli richiesti dalla tecnica esaustiva, quella cioè che esegue la prova di somiglianza tra il blocco di riferimento e il blocco traslato in tutte le posizioni possibili.

Nel presente lavoro non viene trattata la «stima» del movimento, ma si considera solo la «compensazione» del movimento, in inglese detta «motion compensation» e indicata con «MC», intesa come l'insieme delle elaborazioni volte alla specifica utilizzazione delle grandezze caratterizzanti il movimento, tipicamente i vettori movimento, e alla co-decodifica delle stesse per la trasmissione dell'informazione ad esse associata.

Nonostante infatti che la prima sia oggetto di gran parte degli studi pubblicati sull'argomento e rappresenti un formidabile problema di elaborazione del segnale, la prossima comparsa di dispositivi ad alta velocità che renderanno possibile la tecnica di ricerca esaustiva nel prossimo futuro anche per applicazioni televisive, un esempio è riportato in bibliografia 20, e soprattutto la necessità di ridurre al minimo la quantità di informazione associata all'informazione di movimento, coerentemente con l'attività di studio per la riduzione della ridondanza del segnale televisivo che è il tema principale del progetto EU 256, ha spinto ad analizzare più approfonditamente, con lavoro affatto originale, le tecniche di compensazione del movimento e di codifica dell'informazione associata.

La trasmissione dell'informazione relativa al movimento, se non opportunamente codificata, richiede una velocità di cifra troppo alta per le applicazioni di riduzione della ridondanza; per ridurre tale velocità sono stati escogitati diversi metodi, tra cui tecniche basate su menù e tecniche differenziali nel tempo e nello spazio.

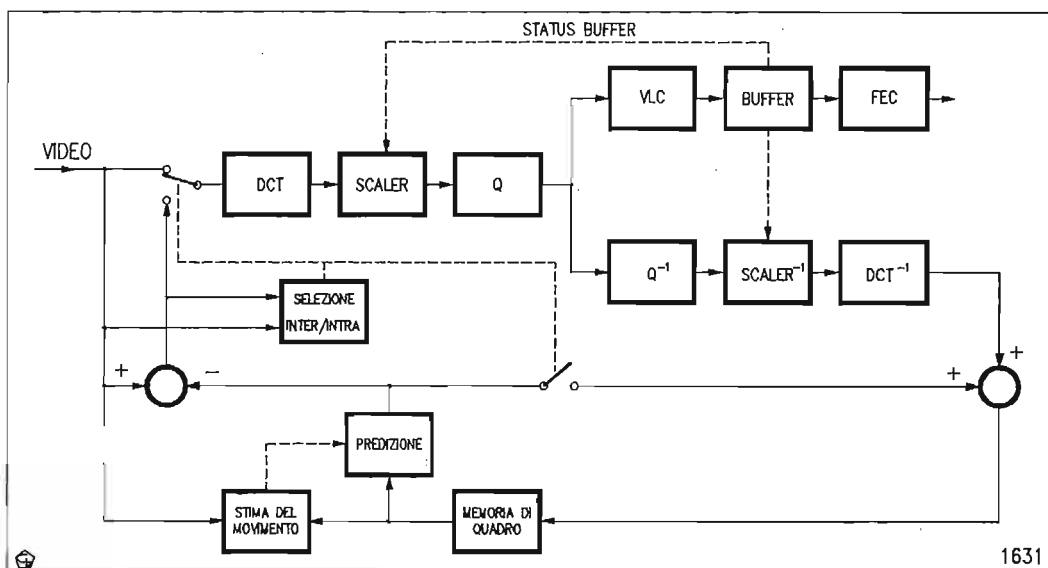
In questo lavoro sono analizzati i due metodi proposti in sede di normativa internazionale riguardo ai codec per applicazioni di qualità contributo per il terzo livello gerarchico CCITT (Mbit/s 34 ÷ 45) (bibl. 2, 3 e 4) e viene proposto un nuovo metodo di codifica dei vettori movimento basato su tecnica differenziale a base statistica, che si è dimostrato insieme molto efficiente e robusto rispetto agli errori di canale.

2. Descrizione del sistema di codifica sviluppato nell'ambito del progetto europeo EU 256

La figura 2 illustra lo schema a blocchi dell'algoritmo di codifica sviluppato nell'ambito del progetto europeo EU 256 per la codifica del segnale televisivo numerico in componenti detto in formato 4:2:2 e descritto nella Raccomandazione n. 601 del CCIR (bibl. 1).

Approfondite descrizioni dell'algoritmo sono reperibili in bibliografia (bibl. 21 e 22), mentre esso viene qui descritto brevemente in funzione di quanto serve per l'argomento trattato.

Per ridurre la correlazione spaziale del segnale televi-

Fig. 2 — Schema a blocchi dell'algoritmo di codifica sviluppato nell'ambito del progetto europeo EUREKA EU 256 (Bit-Rate Reduction System for HDTV Digital Transmission).

sivo viene applicata la trasformata coseno discreta (DCT) bidimensionale su blocchetti di elementi di dimensioni 8 · 8 (8 campioni, 8 linee televisive), ottenendo una matrice di coefficienti DCT di dimensioni identiche.

Nel modo di elaborazione diretto, chiamato «intrafield», gli elementi considerati sono i campioni appartenenti al semiquadro televisivo attuale, quello cioè in corso di elaborazione, ricavati sia dal segnale di luminanza (Y) che dai due segnali di crominanza (CR e CB).

I coefficienti DCT vengono serializzati scandendo in maniera opportuna la matrice degli stessi, di modo che, sia per le caratteristiche della trasformata adottata, sia per le caratteristiche di frequenza tipiche del segnale televisivo, si ottengano sequenze di coefficienti di valore nullo di lunghezza significativa.

Quindi una codifica con parole di lunghezza variabile viene applicata sia ai coefficienti che alle sequenze di valori nulli riducendo la ridondanza statistica del segnale.

Per ridurre una ridondanza temporale si applica la DCT ad elementi ricavati come differenze tra i campioni appartenenti al semiquadro attuale e i campioni appartenenti al semiquadro precedente (modo di elaborazione «interfield») o al semiquadro precedente a distanza di quadro (modo di elaborazione «interframe»).

Nell'algoritmo adottato la scelta del tipo di elementi da elaborare, ossia i campioni del segnale o le differenze tra gli stessi, viene fatta «a priori», cioè prima della trasformata, da un apposito blocco elaborativo.

Questo calcola l'energia degli elementi costituenti il blocco da trasformare nei tre modi suaccennati e seleziona il modo di elaborazione associato al valore minimo dell'energia stessa.

Si assume, e tale assunzione è generalmente verificata in pratica, che il modo di elaborazione così prescelto produca la minore quantità di informazione codificata da trasmettere.

Coerentemente con tale posizione, si assume che qualsiasi tecnica che permetta di minimizzare le differenze tra i blocchi di campioni interessati dalle operazioni dei modi «inter», o, meglio, l'energia delle stesse, porti ad una riduzione della quantità di informazione codificata da trasmettere.

Come esempio, si consideri una ME ideale che permetta di determinare, nel semiquadro precedente stabilito, il blocco esattamente corrispondente a quello in considerazione appartenente al semiquadro attuale, ottenendo così una «compensazione» del movimento perfetta, riconoscibile dall'annullamento delle differenze tra i campioni appartenenti ai due suddetti blocchi: è evidente che la quantità di informazione da trasmettere in questo caso sia la minima in assoluto.

In pratica non si arriva a tale risultato: le parti di immagini non fisse possono subire movimenti non puramente traslatori, ossia roto-traslatori eventualmente con cambiamenti di geometria, che non possono essere convenientemente manipolati dagli algoritmi di ME ordinariamente considerati; i fenomeni di mascheramenti o smascheramento dello sfondo o di altre parti di immagine vicine ai bordi delle parti di immagine in movimento (per i quali il concetto stesso di «movimento» come sopra espresso risulta privo di senso) non possono neppure essi essere manipolati dai comuni metodi di ME; il rumore associato a qualsiasi segnale reale fa sì che parti di immagini considerate identiche dallo spettatore non vengano invece valutate tali dall'elaboratore adibito alla stima del movimento a causa di un eccessivo valore dell'energia degli elementi ricavati dalle differenza tra blocchi dovuti al rumore stesso; inoltre, le elaborazioni del tipo di oggetto sono eseguite nel dominio numerico da processori numerici che presentano una precisione di calcolo finita e non sono in grado di stimare e compensare i movimenti con continuità, ma, anzi, li discretizzano con una certa «risoluzione», tipicamente dell'ordine della metà della distanza tra campioni adiacenti.

Tuttavia, nonostante le limitazioni anzidette, l'applicazione della MC permette, in generale, un miglioramento sensibile della qualità del segnale codificato e una riduzione consistente della velocità di cifra necessaria per la trasmissione dello stesso, come si vedrà nei punti successivi.

Pertanto essa è stata adottata per l'algoritmo sviluppato nell'ambito del progetto EU 256 ed è stata inclusa nelle specifiche sviluppate in sede di normativa internazionale (bibl. 2, 3 e 4) per la definizione delle specifiche per i codec a 34 ÷ 45 Mbit/s e ne è prevista l'adozione nei sisemi per la codifica del segnale numerico HDTV.

Avendo l'algoritmo di codifica sopra descritto la possibilità di utilizzare sia il modo di elaborazione «interfield» che il modo di elaborazione «interframe» sarebbe logico aspettarsi che le migliori prestazioni si possano ottenere applicando la MC ad ambedue i modi di elaborazione.

Le prove fatte hanno dimostrato che per quanto riguarda la quantità di informazione associata al segnale codificato ciò risulta vero, ma considerazioni di tipo realizzativo, in particolare perché viene richiesta una maggiore capacità di calcolo, nonché di tipo trasmissivo, verrebbe in tal caso richiesta una maggiore capacità di canale per poter trasmettere le informazioni di movimento associata ad ambedue i modi di elaborazione sottraendola perciò al segnale televisivo codificato, hanno protato alla decisione di applicare la MC solo al modo di elaborazione «interframe», e ciò sia nell'ambito del progetto EU 256 che in sede di normativa internazionale (bibl. 2, 3 e 4).

In altre parole, gli elementi da elaborare secondo quest'ultimo modo di elaborazione vengono calcolati come sopra descritto permettendo che il blocco appartenente al semiquadro precedente a distanza di quadro (cui appartengono i campioni con cui calcolare le differenze) sia fisicamente traslato della quantità indicata dall'associato vettore movimento e non debba essere coposizionato al blocco di riferimento del semiquadro attuale, come sarebbe invece previsto da un'elaborazione di tipo «interframe» senza compensazione del movimento.

Nelle sedi suindicate si è anche deciso di valutare gli spostamenti con una risoluzione pari a metà della distanza tra campioni di luminanza appartenenti alla stessa linea televisiva, normalmente detta: «risoluzione di mezzo pixel».

Come appare da studi eseguiti presso il *Centro Ricerche* RAI, questo valore di risoluzione viene generalmente considerato un buon valore di compromesso tra l'efficienza ottenibile dall'applicazione di tecniche di compensazione del movimento (si è accennato sopra che una risoluzione finita è possibile fonte di inefficienza) e la potenza di calcolo richiesta per la stima del movimento, dato che in questo caso si possono utilizzare interpolatori lineari

estremamente semplici.

D'altro canto, nel modo di elaborazione «interfield» si confrontano campioni fisicamente non coposizionati a causa dell'interallacciamento del segnale televisivo e, di conseguenza, gli elementi da elaborare vengono calcolati come differenza tra i campioni del semiquadro attuale e i valori di predizione ottenuti interpolando i campioni appartenenti al semiquadro precedente (bib. 2, 3 e 4).

Esistono tecniche di predizione, quali l'LMS (Least Mean Square) (bibl. 23 e 24), che potrebbero convenientemente essere adottate nel modo «interfield» per calcolare i valori di riferimento relativi al semiquadro precedente, e che presentano capacità di compensazione del movimento intrinseche, particolarmente efficaci nel compensare movimenti di piccola ampiezza, richiedendo una minore capacità di calcolo rispetto ad una qualsiasi tecnica di compensazione del movimento e non necessitando di alcuna capacità di canale perché non richiedono l'invio di alcuna informazione al ricevitore.

Inoltre, come si vedrà in seguito, i movimenti più probabili sono di piccola ampiezza, precisamente dell'ordine della distanza tra i campioni del segnale, e pertanto una qualsiasi tecnica di compensazione del movimento applicata al modo «interfield» per essere efficace, necessita, come risulta da prove fatte, di interpolazioni relativamente sofisticate per calcolare valori relativi a frazioni della distanza tra campioni, e queste richiedono notevoli capacità di calcolo in aggiunta a quella già richiesta per l'operazione di prova di somiglianza tra blocchi.

Un parametro importante per dimensionare il sistema è l'ampiezza massima degli spostamenti permessa per la loro compensazione.

Questa, assieme alla risoluzione, influenza sia il numero di bit richiesti per codificare l'informazione di movimento, sia l'area su cui l'apparato che esegue la ME deve estendere la sua ricerca, chiamata «area di ricerca», e quindi la potenza di calcolo necessaria per eseguire la ME, la quale cresce al crescere dell'area di ricerca.

La figura 3 riporta la probabilità di trovare un vettore movimento nell'area di dimensioni indicate: per le due sequenze utilizzate, che si ritengono rappresentative del segnale generato nella normale programmazione televisiva, risulta che più dell'80% dei vettori movimento ha ampiezza tale da essere compreso in una finestra di ± 7 pixel per ± 4 linee; nell'ambito del progetto EU 256 e in sede di normativa internazionale (bibl. 2 e 3) si è adottato una finestra di ± 15,5 pixel per ± 7,5 linee che si ritiene adeguata per il segnale televisivo a definizione normale.

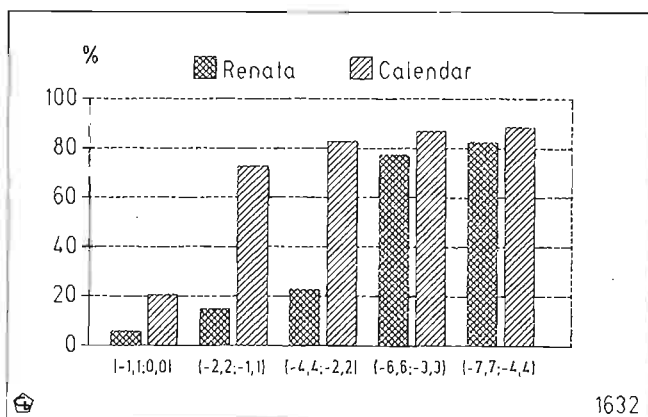



Fig. 3 — Percentuale di vettori movimento compresi nella finestra indicata.

Con queste posizioni si ampiezza massima dei vettori movimento e di risoluzione degli stessi, si richiedono 11 bit per rappresentare dei vettori stessi e la ricerca esaustiva richiede ben 1953 test di somiglianza.

Per limitare il numero di test di somiglianza si sono in passato impiegati metodi chiamati «veloci» che però hanno il difetto di essere meno efficienti della ricerca esaustiva in quanto in certi casi, peraltro frequenti essendo per lo più dovuti al rumore di sorgente, danno un'errata valutazione del movimento.

Chiaramente la scelta del tipo di ME, cioè se adottare la ricerca esaustiva oppure uno dei metodi «veloci», dipende dal costo e dalla realizzabilità del relativo processore, ogni tecnica diversa dalla ricerca esaustiva comportando un aumento della velocità di cifra associata al segnale televisivo codificato o/e un peggioramento della qualità dell'immagine codificata.

Nel seguito viene indicato come sia possibile ridurre la velocità di cifra associata ai vettori movimento.

3. Analisi delle tecniche di compensazione del movimento considerate nell'ambito del progetto europeo EU 256 e in sede di normativa internazionale

I vettori movimento utilizzati dall'algoritmo di codecodifica possono essere il risultato di elaborazioni sui vettori movimento valutati dall'apparato di ME anziché proprio questi ultimi.

In questo lavoro si trascureranno quelle elaborazioni sui vettori movimento volte a determinare dei vettori rappresentativi del movimento reale degli oggetti costituenti l'immagine, in quanto nelle applicazioni previste non necessita tanto il significato fisico del vettore quanto un vettore che porti, come sopra spiegato, ad una riduzione della quantità di informazione associata al segnale televisivo codificato.

Nella realtà infatti, i vettori valutati secondo i due criteri suesposti non necessariamente coincidono.

Inoltre la valutazione dei vettori eseguita secondo il primo criterio, oltre a non essere necessaria, richiede elaborazioni piuttosto sofisticate e più costose che nel secondo caso.

D'altra parte, tale tipo di elaborazione farebbe logicamente parte dell'algoritmo di stima del movimento di cui, come si è detto, non ci si occupa in questo lavoro.

Tuttavia, come si è accennato, nelle applicazioni qui considerate la capacità di canale riservata alla trasmissione dei vettori movimento risulta relativamente piccola e pertanto la quantità di informazione associata ai vettori deve essere ridotta per renderne possibile la trasmissione.

Le due tecniche previste in sede di normativa internazionale (bibl. 4) all'epoca in cui si iniziò il lavoro oggetto di questo articolo, riducono tale quantità di informazione di informazione diminuendo il numero di vettori movimento considerati per la trasmissione stessa.

Il primo metodo suddivide l'immagine televisiva in sottoimmagini, chiamate «zone», e associa un vettore mo-

vimento ad ogni zona; nel seguito si chiamerà tale tecnica: compensazione del movimento «a zone».

Il codificatore e il decodificatore utilizzano, per l'elaborazione degli elementi nel modo di elaborazione interframe, il vettore corrispondente alla zona di immagine in corso di elaborazione, prodotto da un opportuno algoritmo di stima del movimento che opera a livello di zona.

Il numero di vettori movimento trasmessi dipende dal numero delle suddivisioni dell'immagine previste in sede di normalizzazione (in bibliografia 4 viene indicato il numero di 9); nella stessa viene definita anche la precisa locazione nella trama di trasmissione in cui inserire i vettori movimento corrispondenti alle varie zone sicché non si richiede alcuna trasmissione di informazione relativa all'associazione tra zona e vettore essendo essa implicita nel formato della trama di trasmissione.

Il secondo metodo permette la trasmissione di un certo numero prefissato di vettori movimento, i quali vengono inseriti in una lista, chiamata «menu» o «palette», e vengono identificati tramite la loro posizione nella lista; tale tecnica verrà nel seguito indicata come: compensazione del movimento «a menu».

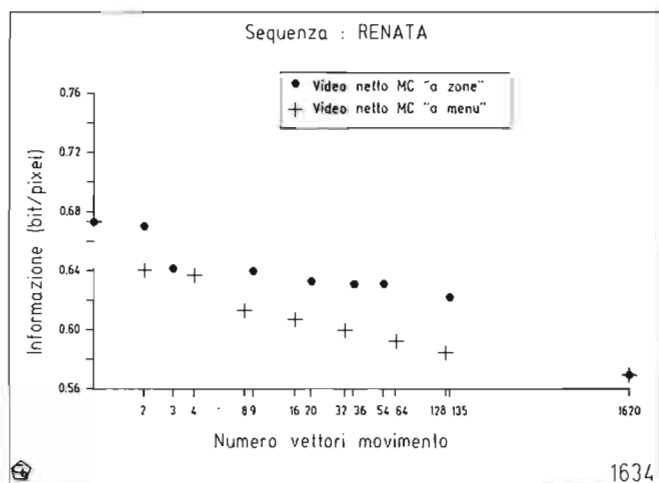
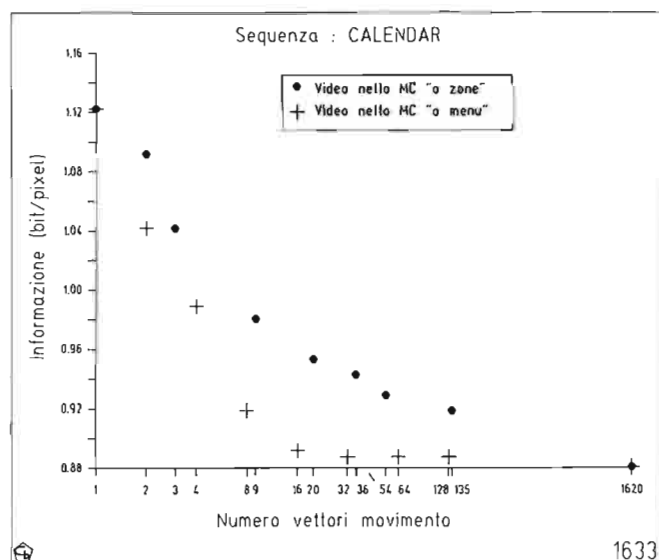



Fig. 4 — Risultati delle analisi comparative tra la tecnica di compensazione del movimento a «zone» e quella «a menu»; valori in bit per pixel.

Per ogni semiquadro in corso di elaborazione vengono scelti i vettori costituenti il menu secondo appropriati criteri, quali per esempio la loro frequenza statistica, e ad ogni blocco viene associato uno dei vettori inclusi nel menu, e ciò sempre con adeguati criteri quali per esempio la distanza vettoriale tra l'associando vettore del menu e il vettore valutato per il blocco in oggetto.

Il numero di vettori movimento trasmessi con questa tecnica dipende dal numero di vettori costituenti il menu stabilito in sede di normalizzazione (in bibliografia 4 viene indicato il numero di 32); contrariamente alla tecnica precedente, la tecnica di compensazione del movimento «a menu» richiede la trasmissione dell'informazione relativa all'associazione tra il blocco di elementi che viene elaborato in modo interframe e il relativo vettore movimento incluso nel menu.

La seconda tecnica richiede quindi una certa capacità di canale per la trasmissione dell'informazione relativa al movimento che in genere viene compensata da una maggiore efficienza di codifica del segnale televisivo.

Per ambedue le tecniche l'efficienza della compensazione del movimento per quanto riguarda la codifica del segnale televisivo aumenta con l'aumentare dei vettori movimento considerati, ossia all'aumento delle zone nella prima tecnica, all'aumentare del numero di vettori costituenti il menu della seconda, raggiungendo il valore massimo nel caso in cui si utilizzi un numero di vettori corrispondenti alla totalità dei vettori stessi.

In quest'ultima configurazione, le zone sono ridotte al blocco base di elaborazione, e la lunghezza del menu permette la trasmissione di tutti i vettori e quindi permette che ad ogni blocco venga associato il proprio vettore.

Le figure 4, 5 e 6 riportano i risultati ottenuti nel corso del lavoro di analisi svolto nell'ambito del progetto europeo EU 256, per ottenere i quali si è simulato l'algoritmo di codifica per la riduzione della ridondanza ivi sviluppato forzandolo a lavorare a qualità dell'immagine costante.

Con tale posizione è stato possibile confrontare l'efficienza delle due tecniche di compensazione sopra descritte sulla base della quantità di informazione associata al segnale televisivo codificato e quantificata in bit per pixel (misura tipica nel campo della riduzione della ridondanza).

Le figure 5a e 5b riportano la riduzione percentuale «D» della quantità di informazione corrispondente al segnale televisivo codificato relativamente al caso in cui non venga applicata alcuna tecnica di compensazione; in altre parole viene presentata la grandezza:

$$D = \frac{I - I_{MC}}{I} \quad [1]$$

dove: «I» rappresenta la quantità di informazione associata al segnale televisivo codificato senza l'applicazione di alcuna tecnica di compensazione del movimento, «I_{MC}» rappresenta la quantità di informazione associata al segnale televisivo codificato con l'applicazione della tecnica di compensazione del movimento specificata.

Tali figure vengono riportate per migliore comprensione dei risultati poiché viene quantificato il beneficio sull'efficienza di codifica del segnale televisivo derivante dalla compensazione del movimento. Si nota che già utilizzando un solo vettore, a volte chiamato «vettore globale» poi-

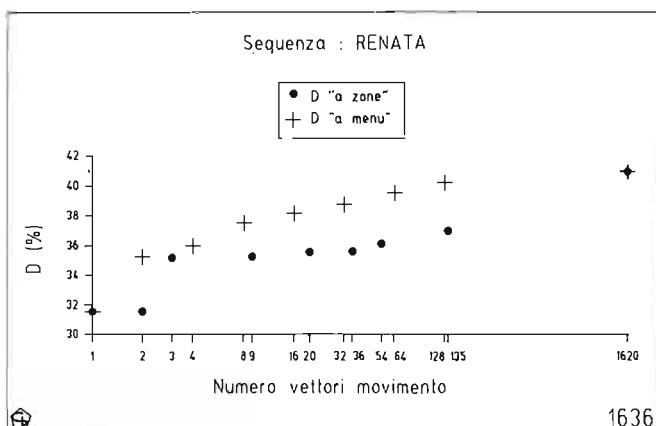
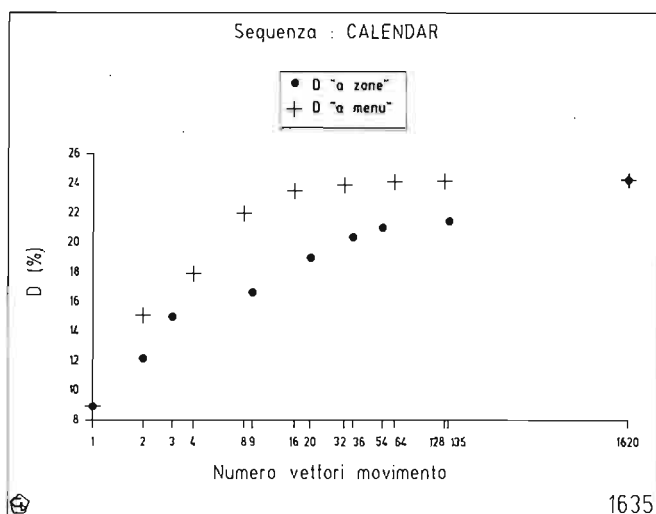



Fig. 5 — Riduzione percentuale «D» della quantità di informazione corrispondente al segnale televisivo codificato relativamente al caso in cui non venga applicata alcuna tecnica di compensazione; valori in percentuale.

ché si riferisce a tutta l'immagine, si ottiene una significativa riduzione dell'informazione relativa al segnale codificato: 31% circa per la sequenza RENATA e 9% circa per la sequenza CALENDAR, nelle simulazioni fatte.

Dai grafici emerge che buone prestazioni si ottengono con le proposte fatte in sede di normativa internazionale (bibl. 4), ottenendo: per la tecnica a zone con 9 zone, valori di «D» pari a 35% e 17% per le sequenze RENATA e CALENDAR rispettivamente, e valori di «D» pari a 39% e 24% per le sequenze RENATA e CALENDAR, rispettivamente, per la tecnica a menu con 32 vettori.

Si noti che utilizzando tutti i vettori a disposizione, nel caso dell'algoritmo sviluppato in EU 256 essi sono 1620 per il segnale in formato 4:2:2, si ottengono valori di «D» pari a 40% e 25% per le sequenze RENATA e CALENDAR rispettivamente.

4. Metodo di codifica dei vettori movimento

Le figure 6a e 6b riportano l'aumento percentuale «A» della quantità di informazione associata al segnale televisivo codificato rispetto al caso in cui vengono utilizzati tutti i vettori; in particolare, «A» viene calcolato come segue:

$$A = \frac{I_{MC} - I_{ott}}{I_{ott}} \qquad [2]$$

dove: «I_{MC}» rappresenta la quantità di informazione associata al segnale televisivo codificato con l'applicazione della tecnica di compensazione del movimento specificata e «I_{OTT}» rappresenta la stessa quantità ottenuta utilizzando tutti i vettori.

I dati riportati nelle figure 6a e 6b evidenziano come l'utilizzazione di un insieme limitato di vettori movimento non permetta di minimizzare l'informazione associata al segnale televisivo codificato, risultando quindi in una perdita di efficienza della codifica del segnale televisivo rispetto a quella ottenibile con l'utilizzazione di tutti i vettori movimento.

È questa considerazione che ha portato a sviluppare una tecnica basata, anziché sulla riduzione del numero di vettori, su di un metodo di codifica degli stessi che permettesse l'invio al decodificatore di tutti i vettori movimento previsti dalla suddivisione in blocchi dell'immagine.

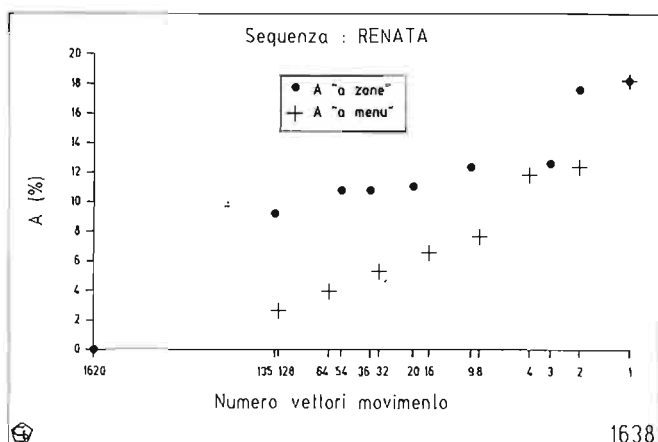
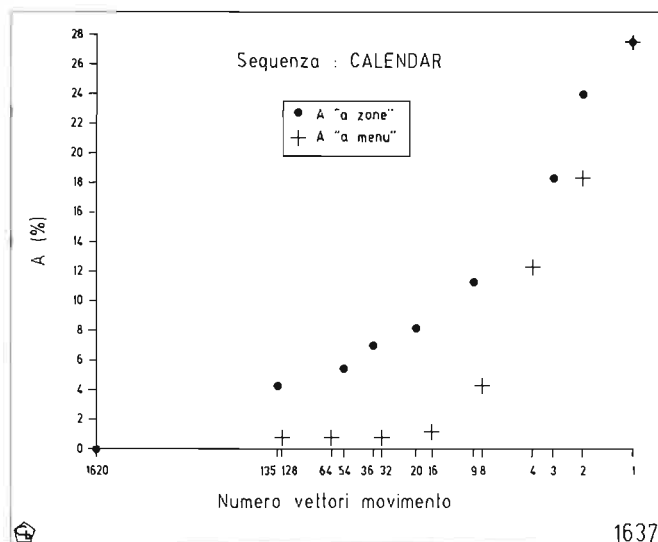



Fig. 6 — Aumento percentuale «A» della quantità di informazione associata al segnale televisivo codificato rispetto al caso in cui vengono utilizzati tutti i vettori; valori in percentuale.

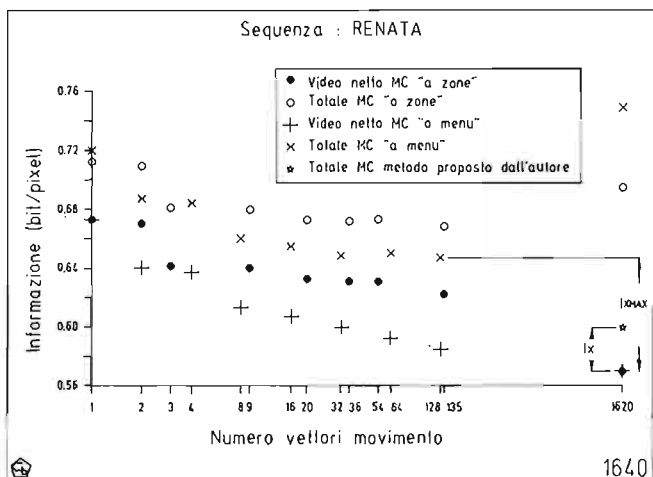
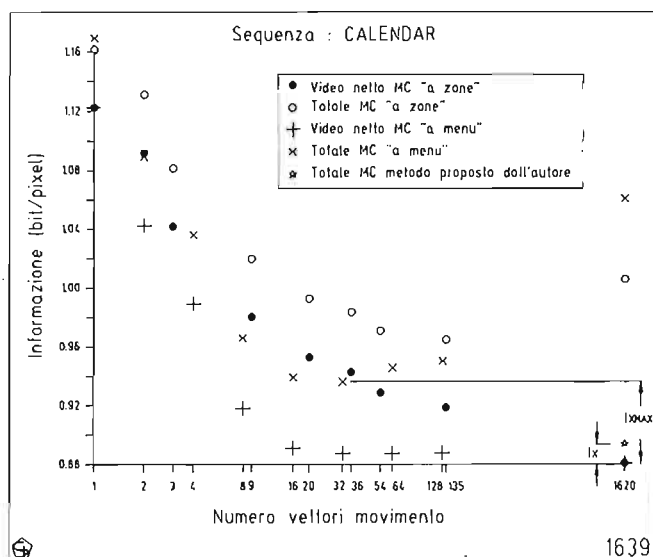



Fig. 7 — Informazione totale e informazione relativa al segnale televisivo codificato per la tecnica di compensazione del movimento «a zone», quella «a menu» e per il metodo di codifica dei vettori movimento descritto nel testo; valori in bit per pixel.

La valutazione completa di una tecnica di compensazione del movimento non può prescindere dalla considerazione dell'informazione associata in qualche modo al movimento stesso che si deve inviare al decodificatore perché questo possa corretamente decodificare i dati che gli giungono; tale informazione verrà nel seguito indicata: informazione «laterale».

Per l'algoritmo di codifica considerato in questo lavoro, tale informazione consiste sia nell'informazione associata ai vettori movimento, sia in quella relativa all'indicazione del modo di elaborazione associato al blocco da elaborare, in quanto questa indica se tale blocco sia da elaborare nel modo interframe con compensazione del movimento oppure in altro modo, ossia, in altre parle, indica se il decodificatore deve applicare la tecnica di compensazione oppure no.

Le figure 7a e 7b, in aggiunta ai dati già riportati nelle figure 4a e 4b, qui di nuovo presentati per confronto, riportano anche l'informazione qui chiamata «totale», ottenuta con somma dell'informazione associata al segnale televisivo codificato e dell'informazione laterale, relativa alle due tecniche precedentemente illustrate e alle due sequenze utilizzate nel corso degli studi.

La valutazione dell'informazione laterale è stata eseguita secondo lo schema illustrato in figura 8 e si sono utilizzati i valori proposti in sede di normativa internazionale (bibl. 4), espandendo opportunamente i concetti ivi illustrati per adeguarsi alle diverse configurazioni considerate.

Dall'analisi delle figure 7a e 7b emerge chiaramente come le tecniche suesposte non possano essere utilizzate

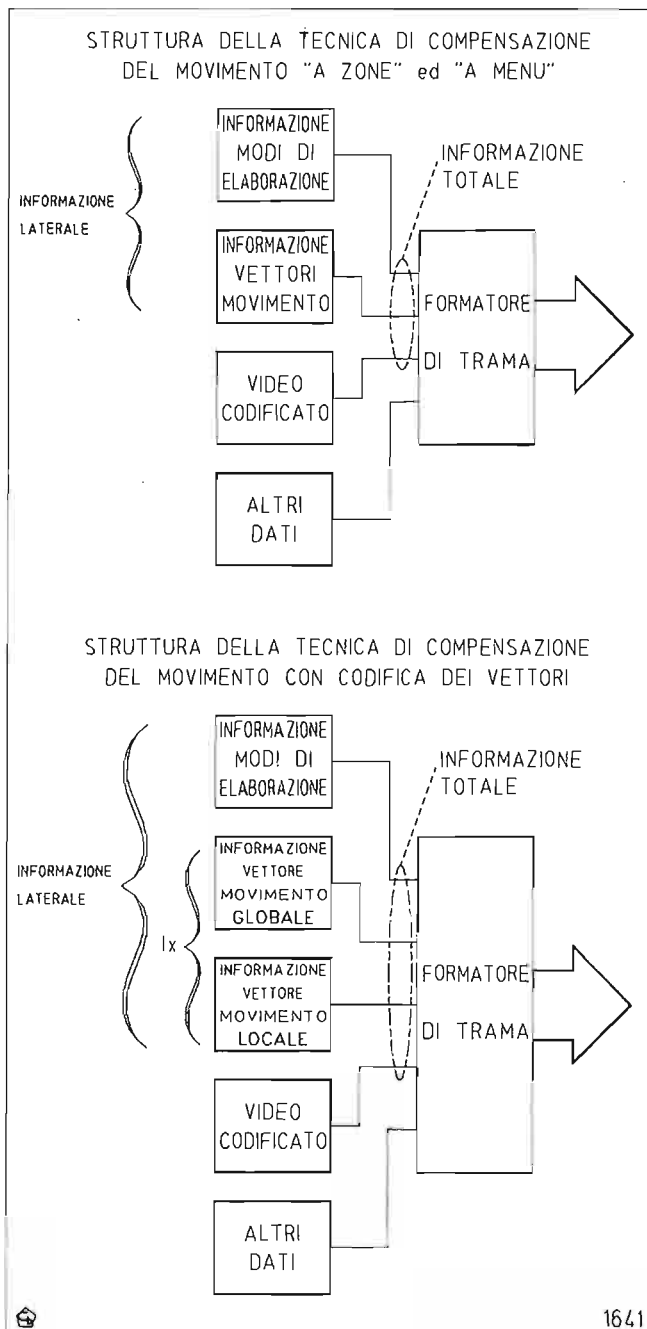



Fig. 8 — Schema illustrativo della procedura di calcolo dell'informazione totale nei casi considerati nel testo. La struttura della tecnica di compensazione del movimento «A ZONE» è uguale a quella «A MENU», ma cambiano le informazioni da elaborare.

convenientemente per l'invio di tutti i vettori movimento, come peraltro era già stato accennato in precedenza.

Emerge anche che per ogni sequenza si trova un valore di minimo dell'informazione totale che un qualsiasi metodo di codifica dei vettori movimento non deve superare al fine di essere considerato maggiormente efficiente.

Di conseguenza, con riferimento allo schema di figura 8 e in accordo a quanto scritto sopra, la quantità di informazione «I_x» associata ai vettori movimento codificati non deve superare un certo valore, indicato nelle figure 7a e 7b con «I_{xMAX}», che non è determinabile a priori, ma è invece dipendente dall'immagine in corso di elaborazione, dall'applicazione specifica (capacità di canale differenti) e dall'implementazione pratica.

Quanto sopra fissa pertanto solo il criterio per la valutazione del metodo di codifica, lasciando una certa discrezionalità nella specifica realizzazione dello stesso.

Il metodo di codifica dei vettori movimento sviluppato presso il *Centro Ricerche* RAI viene qui descritto nelle sue linee generali e, inoltre, ne viene esaminata una particolare realizzazione che è risultata la più efficiente tra quelle prese in considerazione, che sono solo alcune tra le diverse possibili.

Il metodo è schematizzato in figura 9. Esso prevede la valutazione di un vettore che viene chiamato vettore «globale», rappresenttivo, secondo certi criteri quali per esempio la maggior frequenza statistica, dell'intero semiquadro televisivo in corso di codifica e che viene codificato con parole a lunghezza fissa in modo tale da facilitarne la protezione contro gli errori di canale.

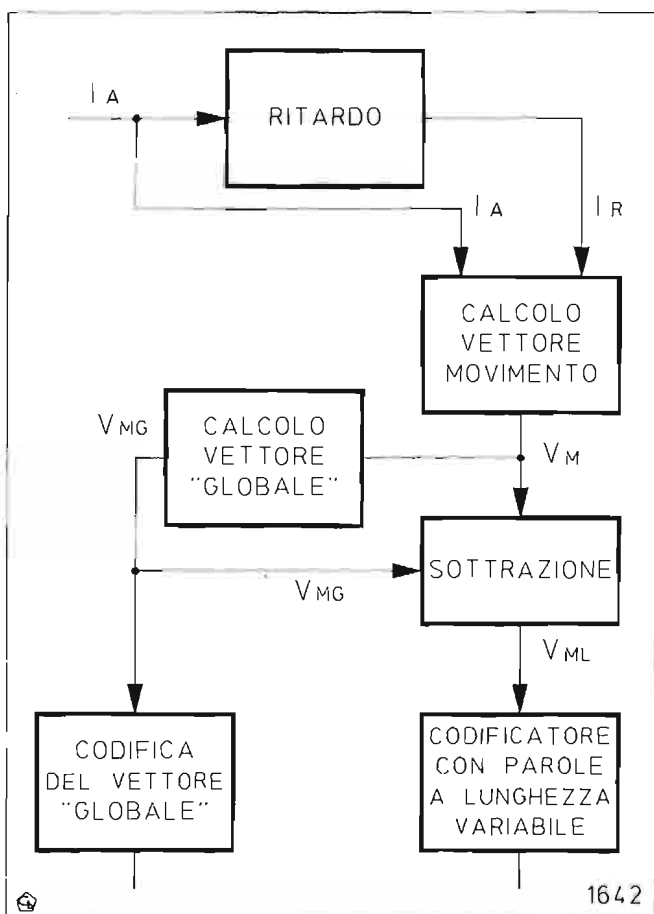



Fig. 9 — Schema del metodo di codifica dei vettori movimento.

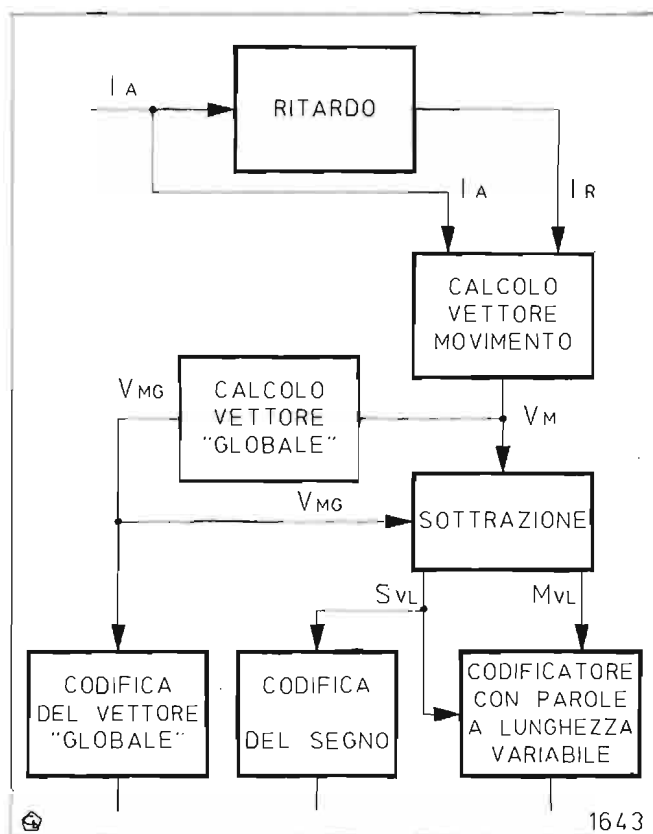



Fig. 10 — Schema di una realizzazione del metodo di codifica dei vettori movimento.

Per ogni blocco di codifica viene associato un vettore, chiamato vettore «locale», valutato come differenza vettoriale tra il vettore globale e il vettore movimento relativo al blocco in oggetto.

La distribuzione statistica delle componenti dei vettori locali presenta un evidente picco nei dintorni dello zero, sicché queste possono essere convenientemente codificate con un adeguato codice a lunghezza variabile (VLC), ottenendo globalmente una consistente riduzione della quantità di informazione associata ai vettori movimento.

Il metodo di codifica risulta robusto rispetto agli errori di canale poiché il vettore globale può essere fortemente protetto contro gli errori di canale grazie alla ridottissima quantità di informazione ad esso associata, e un eventuale errore su un vettore locale porta all'errata decodifica di un solo blocco di elementi del segnale televisivo codificato, evitando così ogni forma di trascinamento dell'errore stesso che è invece caratteristica tipica dei classici metodi di codifica differenziale nel tempo e nello spazio.

La particolare realizzazione di tale metodo che si è dimostrata la più efficiente tra quelle considerate nel corso dello studio viene descritta in figura 10 e prevede di:

— valutare il vettore globale come quello di maggior frequenza statistica nel semiquadro in oggetto,
— valutare i vettori locali come sopra descritto,
— desumere, dall'analisi delle componenti dei vettori locali, se il vettore locale è nullo o, nel caso contrario,

l'indicazione del quadrante in cui il vettore cade, ottenendo un totale di 5 differenti combinazioni;

— inserire le 5 combinazioni di cui sopra nell'informazione di indicazione nel modo di elaborazione, questa codifica essendo altrimenti in accordo con quanto riportato in (bibl. 4) versione B,
— calcolare i valori assoluti delle componenti dei vettori locali,
— codificare vettorialmente i valori assoluti delle componenti con codice a lunghezza variabile «B2-code» se gli stessi sono ambedue diversi da zero oppure codificare scalarmente il valore assoluto della componente orizzontale con codice «B2-code» qualora la componente verticale sia nulla.

Nelle figure 7a e 7b viene anche riportato a titolo di confronto la quantità di informazione totale ottenuta con il particolare metodo qui sopra descritto e si può notare come esso risulti maggiormente efficiente rispetto alle altre tecniche qui considerate per ambedue le sequenze utilizzate in quanto la quantità di informazione «I_x» associata ai vettori movimento risulta in ambo i casi inferiore alla quantità «I_{xMAX}».

5. Conclusioni

Al giorno d'oggi la compensazione del movimento è considerata parte integrante dei sistemi di codifica per la riduzione della ridondanza del segnale televisivo sia a definizione normale che ad alta definizione.

Nell'ambito del progetto europeo EU 256, cui partecipa attivamente il *Centro Ricerche* RAI si è ritenuto di analizzare le tecniche di compensazione del movimento al tempo proposte in sede internazionale (inizio 1989) per verificarne l'efficienza.

In questo lavoro sono stati illustrati i risultati salienti di tale analisi che come risultato ha portato alla considerazione di modificare radicalmente la strategia di codifica delle suddette tecniche e quindi allo sviluppo di un metodo di codifica dei vettori movimento che è risultato insieme efficiente e robusto rispetto agli errori di canale.

Si è data quindi la descrizione del suddetto metodo di codifica in generale e si è illustrata più approfonditamente una delle possibili realizzazioni di cui si sono anche riportati i risultati ottenuti che ne dimostrano l'efficienza in termini di riduzione della quantità di informazione generata.

La robustezza del metodo di codifica rispetto agli errori di canale risulta notevole grazie al tipo di informazione generata dall'applicazione del metodo stesso; in particolare viene minimizzato l'effetto di trascinamento dell'errore tipico delle tecniche di codifica differenziali nel tempo e nello spazio.

Inoltre il metodo richiede una capacità di calcolo per la valutazione delle grandezze da codificare molto ridotta rispetto ad altri metodi quali, per esempio, la tecnica di compensazione «a menu».

Complessivamente quindi si ritiene che il metodo proposto sia particolarmente adatto per la sua adozione in applicazioni di codifica del segnale televisivo e HDTV tramite tecniche di riduzione della ridondanza.

BIBLIOGRAFIA

1. - CCIR - Raccomandazione n. 601-F.
2. - CMTT - Documento n. 303-E, Ottobre 1989.
3. - CMTT/2 - Documento n. 171 (Rev. 1)-E, Marzo 1990.
4. - CMTT/2 - Documento n. 66, Luglio 1988.
5. - LIMB J.O., MURPHY J. A.: *Measuring the Speed of Moving Objects from Television Signals* «IEEE Trans. on Communications», Vol. COM-23, Apr. 1975.
6. - CAFFORIO C., ROCCA.: *Methods for Measuring Small Displacements of Television Images* «IEEE Trans. on Information Theory», Vol. IT-22, No. 5, Sept. 1975.
7. - STULLER J. A., NETRAVALI A. N.: *Transform Domain Motion Estimation* «The Bell System Technical Journal», Vol. 58, No. 7, Sept. 1979.
8. - NETRAVALI A. N., ROBBINS J. D.: *Motion-Compensated Coding: Some New Results* «The Bell System Technical Journal», Vol. 59, No. 9, Nov. 1980.
9. - BOWLING C. D., JONES R. A.: *Motion Compensated Image Coding with a Combined Maximum A Posteriori and Regression Algorithm* «IEEE Trans. on Communications», Vol. COM-33, No. 8, Aug. 1985.
10. - KWATRA C., LIN C., WHYTE W. A.: *An Adaptive Algorithm for Motion Compensated Color Image Coding* «IEEE Trans. on Communications», Vol. COM-35, No. 7, July 1987.
11. - REUTER T., HOHNE H. D.: *Motion Vector Estimation for Improved Standards Conversion* «2nd International Workshop on Signal Processing of HDTV», L'Aquila, Italy, March 1988
12. - GOTZE M.: *Generation of Motion Vector Fields for Motion Compensated Interpolation of HDTV Signals* «2nd International Workshop on Signal Processing of HDTV», L'Aquila, Italy, March 1988.
13. - KERDRANVAT M.: *Hierarchical Motion Estimation and Motion Information Encoding* «3nd International Workshop on HDTV», Torino, Italy, Sept. 1989.
14. - DABNER S. C.: *Real Time Motion Vector Measurement Hardware* «3nd International Workshop on HDTV», Torino, Italy, Sept. 1989.
15. - ROCCA F., ZANOLETTI S.: *Bandwidth Reduction Via Movement Compensation on a Model of the Random Video Process* «IEEE Trans. on Communications», Ottobre 1972.
16. - BROFFERIO S., CAFFORIO C., DEL RE P., QUAGLIA G., RACCIU A., ROCCA F.: *Redudancy Reduction of Video Signals Using Movement Compensation* «Alta Frequenza». Vol. XLIII, N. 10, Ottobre 1974.
17. - NETRAVALI A. N., STULLER J. A.: *Motion-Compensated Transform Coding* «The Bell System Technical Journal», Vol. 58, No. 7, Sept. 1979.
18. - NINOMIYA Y., OHTSUKA Y.: *A Motion-Compensated Interframe Coding Scheme for Television Pictures* «IEEE Trans. on Communications», Vol. COM-30, No. 1, Jan. 1982.
19. - MUSMANN H. G., PIRSCH P.: *Advances in Picture Coding* «Proc. of the IEEE», Vol. 73, No.4, Apr. 1985.
20. - ROTH G., THOLIN P., WIREN B.: *A VLSI for Motion Compensation* «Picture Coding Symposium PCS 87», Stockholm, Sweden, 1987.
21. - BARBERO M., STROPPIANA M.: *Codifica del segnale televisivo numerico: uso della trasformata coseno discreta* «Elettronica e Telecomunicazioni», N. 1, Apr. 1989.
22. - BARBERO M., STROPPIANA M.: *Digital Coding of HDTV based on Discrete Cosine Transform* «Technical Symposium ITU-COM '89», Geneva CH, October 1989.
23. - WIDROW B.M MCCOOL J. M., LARIMORE M. G., JOHNSON JR C.R.: *Stationary and NonStationary learning Characteristics of the LMS Adaptive Filter* «Proc. of the IEEE», Vol. 64, No. 8, Aug. 1976.

LE RIPRESE TELEVISIVE IN ALTA DEFINIZIONE IN OCCASIONE DEI CAMPIONATI MONDIALI DI CALCIO 1990

ROBERTO CECATTO*

SOMMARIO — Il numero «Speciale ITALIA '90» di Elettronica e Telecomunicazioni è stato dedicato in modo particolare agli esperimenti di trasmissione digitale punto-multipunto di televisione ad Alta Definizione affettuati dalla RAI in occasione dei recenti Campionati Mondiali di Calcio. A completamento di quanto trattato nel suddetto numero speciale, si è ritenuto utile descrivere qui le problematiche ed i risultati conseguiti, durante la sperimentazione, in un settore, quale quello delle riprese, che ha contribuito in modo determinante a rendere massima la spettacolarità e il grado di coinvolgimento di chi, in Italia ed in Europa, ha partecipato come spettatore agli esperimenti. L'articolo descrive le fasi ed i vari aspetti, dal progetto di sistema all'addestramento, dell'importante impegno che la RAI ha dovuto affrontare per portare a termine una operazione che oggi si può dire pienamente riuscita, considerata la soddisfazione dimostrata dai partner della RAI in Eureka.

SUMMARY — ***HDTV shootings of football matches during the 1990 World Cup.*** *The issue «Special ITALIA '90» of Elettronica e Telecomunicazioni was dedicated particularly to the experiments of digital HDTV point-to-multipoint transmission carried out by the RAI on the occasion of the FIFA World Cup '90. In addition to the topics dealt with in the special issue, it was deemed useful to discuss also the problems arisen and the results achieved during the experiments in the field of the programme production. This field, in fact, largely contributed to the spectacular quality of the event and to the involvement of everyone, who had the opportunity to watch the programme in HDTV in Italy and in Europe. The article describes the different phases and aspects, from the system project to the training, of the important engagement that the RAI had to face in order to carry out an operation, which now can be regarded as fully successful, considering the consent shown by the RAI's partners in Eureka.*

1. Introduzione

In occasione dei Campionati Mondiali di Calcio (Italia '90) la RAI ha effettuato una campagna di riprese in televisione ad Alta Definizione.

Per quanto concerne le partite giocate nello Stadio Olimpico di Roma, le riprese sono state effettuate dalla RAI nel formato HDTV europeo a 1250 righe/50 trame. Le partite giocate negli altri stadi (Milano, Torino, Napoli, Bari) sono state invece riprese, in virtù di un accordo di coproduzione con la RAI, dalla NHK nel formato HDTV giapponese 1125 righe/60 trame; per il suo interno, la NHK inviava in Giappone i segnali in diretta via satellite oppure, in differita, il registrato per via aerea.

I segnali ripresi in entrambi i formati venivano utilizzati dalla RAI per il suo esperimento di trasmissioni punto-multipunto di televisione ad Alta Definizione numerica (si veda a questo proposito: *SPECIALE ITALIA '90*, «Elettronica e Telecomunicazioni», n. 3, 1990).

Il segnale dello stadio di Roma a formato europeo veniva inoltre inviato al codificatore HD-MAC e, via satellite, veniva ricevuto in cinque postazioni riceventi a Milano, Parigi, Londra, Francoforte ed Eindoven.

Nel presente articolo sono trattati solo i problemi relativi alle riprese allo stadio di Roma che ricadevano sotto la diretta responsabilità RAI.

ITALIA '90, l'avvenimento sportivo dell'anno ha costituito il primo vero banco di prova per la televisione Europea in Alta Definizione.

L'obiettivo è stato quello di verificare le potenzialità produttive dell'Alta Definizione nel campo delle riprese di grandi manifestazioni sportive.

È stata effettuata la copertura in diretta dei sei incontri disputati allo stadio Olimpico di Roma con apparati HDTV 1250/50 per offrire al pubblico italiano ed europeo immagini del Campionato Mondiale di Calcio con qualità artistica e tecnica così elevate da raggiungere i massimi livelli di coinvolgimento e spettacolarità.

In vista di questo importante impegno, la RAI, che si è assunta la responsabilità tecnica ed organizzativa dell'operazione, ha condotto insieme ai partner di Eureka un lavoro preparatorio articolato in tre fasi, il progetto sistemistico, la sperimentazione e l'addestramento, la pianificazione e la logistica.

2. Progetto e realizzazione del sistema di produzione

La RAI ha imposto per il sistema di produzione requisiti particolarmente «elevati» tenuto conto che la tec-

* Ing. Roberto Cecatto del Supporto Tecnico della RAI - ROMA.
Dattiloscritto pervenuto alla Redazione il 12 ottobre 1990.

nologia HDTV non ha ancora raggiunto la maturità e la flessibilità di quella convenzionale a cui siamo abituati.

Il risultato di tali richieste è stata la messa a punto di un sistema di produzione mai realizzato prima, per dimensioni e flessibilità, neanche sperimentalmente.

Il raggiungimento di tale ambizioso obiettivo richiedeva la disponibilità di un numero elevato di apparati: 6 ÷ 7 telecamere, 2 ÷ 3 videoregistratori per i replay, titolatrice, una regia adeguata, con un mixer a 12 ÷ 16 ingressi, 2 videoregistratori per la registrazione di due master.

Dopo ampie consultazioni con il Direttorato di Eureka, la disponibilità degli apparati è stata assicurata mediante la concessione dell'uso di due grandi mezzi mobili: uno tedesco di proprietà della BTS, l'altro spagnolo di propriеà della TVE, ognuno dotato di 3 telecamere, 2 videoregistratori, 1 mixer video, 8 ingressi, più vari apparati accessori.

Il sistema di produzione doveva garantire anche la messa in onda contemporanea di due programmi HD diversi per due utenze diverse. Infatti, si doveva alimentare il codificato HD Numerico con un programma «Italiano», mentre con un programma «Eureka», a cui era abbinato un audio a 6 canali e titoli di testa e di coda differenti, si doveva alimentare il codificatore HD MAC.

Inoltre, già da quattro ore prima di ogni partita, i due codificatori dovevano essere alimentati con due differenti programmi di test video ed audio per i circuiti via satellite per i punti di visione.

L'impostazione, in sede di progetto, di caratteristiche di ampia flessibilità del sistema ha consentito il soddisfacimento di tali richieste.

Ai fini di una migliore gestione operativa, la RAI ha proposto la centralizzazione dell'impianto, utilizzando il mezzo mobile della BTS come «Master» e quello della TVE come «slave».

Solo il mezzo mobile tedesco infatti, disponendo di una matrice di commutazione video 10 · 10 e di numerosi distributori video equalizzati, rispondeva alla flessibilità richiesta.

2.1 TELECAMERE

Un altro problema da risolvere è stato l'accoppiamento colorimetrico tra le varie telecamere.

Le telecamere del mezzo BTS sono dotate di tubi Plumbicon da 1 pollice e 1/4 mentre le telecamere del mezzo TVE (sempre di fabbricazione BTS) sono dotate di tubi Saticon da 1 pollice.

Era lecito attendersi un comportamento colorimetrico differente, considerando il fatto che le partite iniziavano al tramonto, e quindi si passava dinamicamente da una condizione di luce mista (naturale-artificiale) ad una luce completamente artificiale.

Per questo motivo la RAI, nel progetto, ha indicato la necessità di centralizzare anche i controlli video di tutte le telecamere. Questa esigenza tecnica è stata soddisfatta, posizionando i tre pannelli di controllo delle telecamere TVE accanto ai tre pannelli di controllo delle telecamere BTS e prolungando la linea di collegamento dati (di tipo party line) dai CCU (Camera Control Unit) spagnoli. In questo modo controllando su di un'unica linea monitoria di riferimento (monitor colori HDTV — waveform monitor) i parametri del segnale video di tutte le telecamere, è stato possibile non rendere apprezzabile la differenza di comportamento di tubi di ripresa diversi.

Gli obiettivi erano 2 Fujinon 14 · 20, un Angenieux

Fig. 1 — La figura illustra una delle postazioni allestite per le riprese, con telecamera ad Alta Definizione della BTS allo stadio Olimpico di Roma in occasione dei Campionati Mondiali di Calcio «ITALIA '90».

soluzione. Innanzitutto è stata ottimizzata l'ergonomia dell'operatività consentendo al tecnico che gestiva le due macchine, di operare attraverso dei comandi remoti. In secondo luogo, per avere un riferimento ai fini della ricerca di una sequenza da riproporre «in onda», è stata convertito il segnale HD, di ingresso/uscita per ognuno dei due videoregistratori, in PAL. Su queste immagini PAL è stato possibile intarsiare il codice di tempo in uscita dai videoregistratori. In questo modo quando il video andava sul nero durante l'indietro veloce o l'avanti veloce, permaneva il riferimento temporale sui monitor PAL che ha consentito all'operatore di cercare, per esempio, l'inizio dell'azione di un goal, stimandone la durata complessiva in «x» secondi. In questo modo, grazie anche alla perizia del personale addetto, sono stati riproposti con successo numerosi replay delle azioni più importanti delle partite riprese in HDTV.

2.4 AUDIO

Come è stato accennato all'inizio, si è cercato di realizzare un programma spettacolare e coinvolgente.

Per ottenere questo risultato, insieme alle immagini di elevata qualità in HDTV (quindi formato panoramico, possibilità di proiezione di qualità su grande schermo), è stato offerto al pubblico un audio associato «particolare».

Il punto di partenza è stato l'utilizzazione del programma audio stereo della multilaterale PAL che è stato realizzato dai tecnici della RAI con notevole cura. Infatti per catturare chiaramente i suoni, le grida, i rumori prodotti sul campo, sono stati utilizzati 20 microfoni assai direttivi, collocati perimetralmente al campo di gioco, posizionati ad un'altezza di circa due metri ed orientati leggermente verso il manto erboso, in modo tale da evitare al massimo i rientri del pubblico sugli spalti.

Inoltre, ogni microfono veniva guidato da un operatore, attraverso un manubrio, per sfruttarne al massimo la caratteristica di direttività. Un tecnico audio quindi selezionava continuamente il microfono migliore in quel momento per la riproduzione dei suoni dell'azione di gioco.

Il segnale audio del campo veniva quindi miscelato ad altri segnali che creavano l'ambiente stereo dello stadio, rispetto al fronte di ripresa delle telecamere. Questa è stata la composizione del segnale stereo di base. Su questo segnale stereo è stato quindi miscelato centralmente il giornalista che ha effettuato il commento delle immagini in HDTV.

Oltre a questo canale stereo, è stato anche prodotto un terzo canale che possiamo definire «posteriore», utilizzando dei microfoni ed una testa artificiale posizionati dietro gli spalti della tribuna. Al codificatore HDTV numerico è stato quindi inviato il seguente audio associato:

canale 1 sinistro	Effetto pubblico sx
canale 2 centrale	Effetti campo-cronista italiano
canale 3 destro	Effetti pubblico dx
canale 4 posteriore	Effetti tribuna posteriore centrale

In aggiunta al cronista italiano, sono stati presenti altri tre cronisti: inglese, francese e tedesco, i quali hanno prodotto quattro canali di commento che, unitamente al canale stereo descritto sopra, formavano l'audio associato al segnale video inviato al codificatore HD-MAC:

canale 1	Effetti pubblico sx	Effetti campo
canale 2	Effetti pubblico dx	
canale 3	Commento inglese	
canale 4	Commento francese	
canale 5	Commento tedesco	
canale 6	Commento italiano	

Tutti questi canali audio prodotti, oltre essere stati messi in onda in diretta, sono stati registrati su diversi supporti magnetici sia dedicati, come i registratori audio a bobine da 1/4 di pollice, sia abbinati alla registrazione video (video-registratore HDTV e Betacam).

Fig. 3 — Visione d'insieme delle postazioni attrezzate riservate alla stampa ed ai cronisti radio-televisivi presso lo stadio Olimpico di Roma.

2.5 ALTRI APPARATI

Come nel caso precedentemente descritto relativo ai VTR dedicati ai replay, per soddisfare i requisiti di produzione è stato necessario studiare particolari soluzioni tecniche di cui ora parleremo.

Essendo assolutamente necessario, come si può ben intuire, prevedere un minimo di grafica elettronica (punteggio, formazioni, cronometro, ecc.) e non essendo disponibile sul mercato alcun apparato funzionante nel formato europeo 1250/50, è stata ideata la seguente configurazione. Il down-converter, a partire dal trisync (HD) di sistema, generava con coerenza di fase un sincronismo PAL. Con questo si è sincronizzato un impianto PAL consistente in un generatore di caratteri ed un orologio-cronometrico che combinati in un mixer video convenzionale hanno alimentato un up-converter realizzato a livello di prototipo dalla Seleco. Occorre far presente che il processo di upconvertion, nel caso di caratteri statici, raggiunge una qualità particolarmente buona, in quanto l'interpolazione tra due righe adiacenti per crearne una terza, produce, con immagini fisse, buoni risultati.

In uscita dall'up-converter si è poi usato il segnale di luminanza (Y) che è stato usato come chiave per l'inseritore di titoli del mixer HDTV presente nella regia del mezzo BTS. In questo modo si è potuto produrre, in diretta, una grafica di buona qualità e con una agevole operatività.

Un'altra esigenza da soddisfare è stata quella di fornire ai quattro cronisti del programma HDTV, presenti nella tribuna stampa, le immagini generate dal sistema di produzione per il loro commento.

Essendo costoro situati a circa 400 metri di distanza dai mezzi mobili, è stato necessario usare un down-converter per tramutare il segnale HD in PAL, quindi utilizzare un link a 70 MHz (modulatore/cavo coax/demodulatore) per trasportare il segnale senza perdite di qualità fino alla tribuna stampa. Infine, attraverso un distributore video il segnale è stato inviato ai monitor convenzionali presenti nella postazione standard. Per la conversione è stato usato il formato «lettor box» che consente di visualizzare tutta l'immagine formato 16/9 su un display formato 4/3. Dal punto di vista impiantistico, anche se tutto il sistema è stato cablato in componenti (Y, Cr, Cb) erano presenti degli apparati con ingresso e uscite in R, G, B. Per l'inserimento di questi apparati nell'impianto, è stato necessario utilizzare delle matrici e dematrici di elevata qualità che sono state realizzate dal *Centro Ricerche* RAI.

3. Personale e preparazione logistica

La RAI ha avuto la responsabilità tecnica e logistica della produzione. Tutto il personale tecnico operativo, dal direttore di produzione alla regista, dai cameramn ai tecnici, per un totale di 24 elementi è stato fornito dalla RAI. Erano presenti anche alcuni tecnici della BTS e della TVE per garantire una pronta assistenza in caso di guasti o mal funzionamenti.

Indubbiamente un fattore molto importante per il buon esito di operazioni così delicate è la preparazione logistica. La RAI ha ricevuto un'area di circa 100 metri quadrati prospicente lo stadio, solo 4 settimane dall'inizio della prima partita. In questo periodo sono state portate a termine le opere edili per la realizzazione delle varie infra-

Fig. 4 — Particolare della figura 3. Il cronista delle riprese effettuate in Alta Definizione riceveva su uno dei due monitor le immagini relative al segnale trasmesso, convertite nello standard PAL; sull'altro monitor riceveva le immagini della ripresa convenzionale (PAL), oppure poteva selezionare informazioni relative alle squadre, ai giocatori, ecc.

Fig. 5 — Vista dell'area dedicata all'Alta Definizione televisiva sul piazzale Belvedere, presso lo stadio Olimpico di Roma, in occasione dei Campionati Mondiali di Calcio Italia '90.

strutture necessarie, e sono stati approntati tutti i servizi. Oltre ad una recinzione ed al posizionamento di vari container per ospitare gli apparati, è stato realizzato un tunnel per il passaggio dei cavi da e verso lo stadio sotto una strada a scorrimento veloce e un passaggio aereo di 40 metri per portare direttamente i cavi camera dall'uscita del tunnel alla parte alta della tribuna.

Per quanto riguarda le comunicazioni sono state installate 6 linee telefoniche SIP, di cui due con fax e 6 linee collegate alla rete interna RAI. Erano disponibili inoltre 10 ricetrasmittenti portatili per comunicazioni di emergenza.

Imponente è stata la realizzazione dell'impianto elettrico. È stato approntato un sistema capace di erogare 300 kVA con una stabilizzazione della tensione contenuta entro l'1% grazie all'utilizzo di un UPS. Per garantire la continuità dell'alimentazione erano previste due sorgenti, la rete ENEL e un gruppo elettrogeno, commutate automaticamente dall'UPS. È stata posta particolate attenzione alla realizzazione dell'impianto di terra, configurandolo a stella per evitare possibili interferenze sui segnali video audio e utilizzando cavi di grossa sezione (90 mm quadrati).

4. Addestramento

La necessità di ideare un nuovo stile di ripresa, tenuto conto della novità e potenzialità del mezzo tecnico a disposizione, ha richiesto un addestramento preventivo per tutto il personale tecnico operativo. Due mesi prima del «Campionato del Mondo» sono state effettuate le riprese di 4 partite di cui due giocate allo stadio S. Siro di Milano, provando varie posizioni per le telecamere, gestioni differenti dei replay, diversi modelli di regia.

Al termine del periodo di addestramento, il personale ha conseguito una migliore e più sicura operatività; sono state inoltre acquisite numerose informazioni di carattere tecnico che sono risultate estremamente utili per il progetto definitivo del sistema di produzione.

5. Conclusioni

Sono state prodotte e messa in onda 13 ore di televisione in Alta Definizione «Live». Per quanto concerne gli apparati di produzione operanti nel formato di scansione europeo 1250 righe/50 trame, è stata acquisita da parte della RAI e dei partner del progetto europeo Eureka EU 95, una esperienza produttiva unica.

Oltretutto questa operazione è stata un'anticipazione di due anni rispetto ai grossi impegni pianificati dallo stesso Eureka EU 95 per il 1992, per la copertura di importanti avvenimenti quali le Olimpiadi invernali di Albertville, i giochi olimpici di Barcellona e l'Expo Universale di Siviglia.

L'intera catena di produzione è stata utilizzata e verificata in ogni sua parte, traendone indicazione utili e stimolanti per le industrie.

La riuscita dell'operazione, confermata dalla piena soddisfazione di tutti partner di Eureka EU 95, ha confermato in primo luogo la capacità professionale e produttiva della RAI e, in secondo luogo, ha sicuramente legittimato e dato impulso agli sforzi che broadcasters e industrie stanno compiendo per la diffusione della «Nuova Televisione».

(3938)

Assegnazione del Montreux Achievement Golden Medal all'ing. Marzio Barbero del Centro Ricerche RAI (Montreux, 13-18 Giugno 1991)

Il Simposio Internazionale e l'Esposizione Tecnica che si tengono a Montreux, con cadenza biennale, rappresentano l'evento tecnico più importante, tra quelli che si celebrano in Europa, nel settore della televisione professionale. Alla manifestazione, di portata mondiale, si danno appuntamento gli operatori più qualificati impegnati nello sviluppo di sistemi ed apparati per la produzione e trasmissione televisiva.

In occasione della cerimonia inaugurale viene solitamente assegnata, come risultato di una severa selezione effettuata da un Comitato Internazionale, una medaglia d'oro (The Montreux Achievement Gold Medal), a chi si è particolarmente distinto nello sviluppo tecnologico della televisione.

Quest'anno il riconoscimento è stato attribuito all'ing. Marzio Barbero del *Centro Ricerche* RAI con la seguente motivazione: «per il suo contributo alla trasmissione numerica della TV, includendo la HDTV, basata su tecniche DCT», cioè per il decisivo apporto dato allo sviluppo di tecniche di codifica numerica con riduzione delle ridondanze nei segnali televisivi, sia convenzionali che ad Alta Definizione.

I ricercatori del *Centro Ricerche* RAI hanno contribuito in particolare all'attività del progetto europeo Eureka EU 256 che ha visto impegnati, in Italia, la Telettra S.p.A. e la RAI-Radiotelevisione Italiana ed in Spagna, Retevision, Telettra Espana e l'Universidad Politecnica de Madrid, nello sviluppo dei sistemi e degli apparati televisivi numerici di co-decodifica.

Le tecniche di codifica numerica sono considerate di grande importanza strategica per l'evoluzione dei sistemi televisivi in quanto lasciano prevedere una vera e propria rivoluzione in termini di qualità e varietà dei servizi da inviare all'utente.

A partire dal 1986 si è sviluppata una notevole attività in ambito internazionale per la definizione di sistemi atti a consentire la trasmissione numerica del segnale televisivo attraverso le attuali reti di collegamento. Grazie al contributo italiano è stato possibile, nonostante posizioni notevolmente contrastanti, raggiungere in Europa un primo accordo relativo alla televisione convenzionale, che si rifletterà anche sulla normativa mondiale.

L'attività del progetto EU 256 ha avuto il suo massimo successo con la realizzazione del co-decodificatore numerico per segnali televisivi ad Alta Definizione, che ha consentito di effettuare, in prima mondiale, gli esperimenti di trasmissione di HDTV numerica via satellite punto-multipunto in occasione dei Campionati Mondiali di Calcio giocatisi in Italia nell'estate del 1990 (vedere Elettronica e Telecomunicazioni, n. 3, 1990).

(3934)

CODIFICA DEL SEGNALE TELEVISIVO NUMERICO: STIMA E COMPENSAZIONE DEL MOVIMENTO

S. RAVERA, L. ROSSATO *

SOMMARIO — Presso il Centro Ricerche RAI è oggetto di studio la codifica a riduzione di ridondanza del segnale televisivo numerico, volta a permettere la trasmissione di quest'ultimo su canali di media e bassa capacità. La compressione attuata sul segnale mediante l'eliminazione della ridondanza spazio-temporale risulta particolarmente utile per la trasmissione del segnale HDTV digitale, dove il bit-rate *richiesto per la codifica PCM si aggira intorno ai 900 Mbit/s. Con tecniche basate sull'uso della Trasformata Coseno Discreta, descritte in articoli precedenti, si possono ottenere collegamenti a «qualità contributo» utilizzando un bit-rate di soli 34 Mbit/s per il formato 4:2:2 e di circa 140 Mbit/s per l'HDTV. La «qualità distribuzione» può essere conseguita con un'occupazione di banda grossomodo dimezzata. Volendo ridurre ulteriormente la quantità di informazione da trasmettere, occorre impiegare tecniche più sofisticate di quelle già descritte; in questo articolo si discutono le prestazioni dell'algoritmo di «compensazione del movimento», ponendo un accento particolare sulla fase di stima dei vettori movimento.*

*SUMMARY — **Digital TV signal coding: motion estimation and compensation.** At the RAI Research Centre, studies and optimizations are being carried out on the redundancy reduction coding of the digital TV signal, in order to transmit it on medium or low capacity channels. The compression obtained on the information flow with the elimination of the spatial-temporal redunancy is especially attractive for the transmission of the digital HDTV signal; in this case, the required bit-rate for PCM coding is about 900 Mbit/s. Using techniques based on the Discrete Cosine Transform, which have been described in previous articles, we can obtain «contribution quality»links with a total bit-rate of only 34 Mbit/s for the 4:2:2 format and about 140 Mbit/s for the High Definition TV (HDTV). «Distribution quality» can be reached with 17 and 70 Mbit/s. If further reductions are required on the transmitted information flow, more sophisticated tecniques are to be considered; in this artiche we deal with the performance of the «motion compensation» algorithm, particularly focusing on the motion estimation criteria.*

1. Introduzione

Il co-decodificatore video sviluppato nell'ambito del progetto europeo EU 256, descritto in precedenza su questa stessa rivista (bibl. 10 e 11) opera una riduzione considerevole del flusso di informazione necessario per la trasmissione del segnale video digitale, senza introdurre degradamenti apprezzabili nella qualità dell'immagine.

Ciò è possibile poiché viene eliminata, o fortemente ridotta, la *ridondanza* presente nel segnale, ossia l'informazione che non è necessario trasmettere, in quanto ricostruibile al ricevitore con tecniche statistiche; come si può notare, una tale operazione risulta del tutto trasparente, dato che non viene persa alcuna informazione utile.

La compressione avviene procedendo in due direzioni distinte: si sfrutta sia la correlazione spaziale, mediante l'applicazione della trasformata DCT (bibl. 9), sia quella temporale. Per quest'ultima, l'accorgimento usato consiste nel codificare, invece dei dati video, le differenze tra due quadri o semiquadri contigui (quando ciò, naturalmente, dia luogo ad un flusso di informazione minore).

Ora, in presenza di immagini in movimento, la correlazione temporale viene ad essere fortemente ridotta; ne consegue un aumento del bit-rate necessario alla codifica, ovvero, se il flusso di informazione è limitato dalla capacità del canale di trasmissione, un peggioramento di qualità, conseguente alla necessità di scartare anche parte dell'informazione utile.

Tale problema può essere risolto facendo uso di un'opportuna tecnica di **compensazione del movimento**, in grado di tenere in conto, nella codifica delle differenze interfield o inter-frame, degli spostamenti di tutta l'immagine o di parti di essa.

In particolare, nel caso del codec EU 256, il semiquadro viene diviso in blocchi di 16·8 «pel»; su ognuno di essi viene valutato, mediante un algoritmo noto come *block matching*, un vettore movimento locale, che viene poi utilizzato per ricavare le differenze di predizione temporale. Si suppone che i blocchi siano abbastanza piccoli da possedere solo movimenti di traslazione[1].

In previsione di una realizzazione hardware a breve scadenza, al *Centro Ricerche* RAI furono effettuate numerose prove di simulazione, tese a verificare l'effettiva utilità della compensazione movimento ed a valutarne quantitativamente le prestazioni, in termini sia di riduzione del bit-rate di codifica che di qualità dell'immagine ricostruita al ricevitore. In questo articolo riportiamo i risultati di tale lavoro; inoltre presentiamo una panoramica delle diverse tecniche ad oggi utilizzate per la stima del movimento. Quest'ultima operazione, come si vedrà, rappresenta la parte più critica del processo di compensazione.

---

* Ing. Secondina Ravera, ing. Luca Rossato del Politecnico di Torino.
Dattiloscritto pervenuto alla Redazione il 14 agosto 1991.

1 Naturalmente, questa ipotesi non è sempre verificata. Tuttavia, la compensazione di movimenti diversi da quello traslatorio presenta grandi difficoltà realizzative, e almeno per il momento questa strada non è stata percorsa.

2. Tecniche di stima del movimento

Le tecniche attualmente impiegate per la stima del movimento nell'elaborazione di immagini televisive possono essere suddivise in quattro categorie, in accordo a quanto espresso nella letteratura: esse si basano rispettivamente sull'astrazione dei contorni, su metodi differenziali, sulla trasformata di Fourier o sul confronto diretto di blocchi d'immagine (*block matching*).

2.1 ESTRAZIONE DEI CONTORNI

L'uso di questa tecnica prevede il riconoscimento, all'interno dell'immagine televisiva, di particolari caratteristiche degli oggetti che in essa sono contenuti (bibl. 1-3); in genere si punta soprattutto all'individuazione dei bordi degli oggetti, che sono facilmente isolabili mediante un filtraggio passa-alto dell'intera immagine. Una volta avvenuta l'esaltazione dei contorni, si applica una delle tecniche differenziali descritte nel paragrafo 2.2.

Si ottiene in tal modo un'informazione di movimento relativa a ben determinate zone dell'immagine; in seguito è possibile eseguire delle interpolazioni al fine di ricavare i vettori movimento per le aree non considerate.

Come si può intuire, questo modo di procedere è particolarmente adatto al riconoscimento ed all'inseguimento (*tracking*) degli oggetti presenti in una scena; pertanto si presta, ad esempio, alla ricostruzione di strutture tridimensionali a partire da immagini bidimensionali, oppure come ausilio nei sistemi di visione artificiale.

Per quanto riguarda la codifica a riduzione di ridondanza del segnale televisivo, al contrario, la tecnica descritta non è stata sinora oggetto di grande attenzione. Ciò è dovuto principalmente a due motivi: il primo è la notevole complessità realizzativa, dovuta alla necessità di successivi filtraggi dell'immagine (prima con un passa-alto per esaltare i gradienti di luminanza che costituiscono i contorni, poi con un passa-basso per ridurre l'effetto del rumore di quantizzazione e per aumentare la continuità delle linee di confine tra gli oggetti).

L'altro motivo è intrinseco alla modalità con cui si opera, e riguarda l'impossibilità di determinare il movimento di aree d'immagine senza contorni netti (ad esempio zone uniformi o quasi uniformi) oppure di oggetti con un grande contenuto di alte frequenze spaziali; in questi ultimi, l'eccesso di dettaglio rende molto critica l'individuazione di caratteristiche particolari come i contorni.

Nonostante le limitazioni a cui è soggetta, questa tecnica di stima del movimento appare agli autori suscettibile di applicazioni promettenti. In effetti, operando a bit-rate sempre più ridotti, può nascere la necessità di effettuare dei sottocampionamenti temporali del segnale televisivo; in pratica, vengono codificati solo alcuni semi-quadri (per esempio uno ogni due), ed i rimanenti vengono ricostruiti al ricevitore mediante interpolazioni. Poiché in questo caso è necessario disporre di vettori movimento che diano informazioni il più possibile affidabili sul **reale** movimento degli oggetti della scena, la tecnica di estrazione dei contorni potrebbe avere un'importanza fondamentale, soprattutto se adeguatamente migliorata attraverso ulteriori studi ed ottimizzazioni.

Fig. 1 — Algoritmo di stima del movimento di Limb e Murphy. La linea tratteggiata indica la posizione del fronte nel quadro precedente.

2.2 TECNICHE DIFFERENZIALI

I primi algoritmi per la stima del movimento nell'ambito di immagini televisive sono stati proposti da Limb e Murphy (bibl. 4) nel 1975 e da Cafforio e Rocca (bibl. 5) nel 1976.

L'algoritmo di Limb e Murphy può essere spiegato considerando un fronte di luminanza in movimento, come mostrato in figura 1. Lo spostamento S viene stimato come segue:

$$[1] \qquad \hat{S} = \hat{dx} = \frac{\Sigma_A \mid DQ \mid}{H}$$

dove $|DQ|$ è il modulo del segnale differenza tra due quadri successivi, mentre H rappresenta l'ampiezza del fronte. La sommatoria si intende estesa all'area A, definita dalla zona in cui la differenza $|DQ|$ supera una soglia prestabilita. Il numeratore corrisponde all'area tratteggiata in figura 1, il denominatore all'altezza del parallelogramma.

L'equazione [1], di per sé, non fornisce nessuna informazione sulla direzione dello spostamento; è comunque possibile risalire a quest'ultima in modo abbastanza semplice osservando gli istanti di partenza del fronte nei quadri successivi.

L'approccio di Cafforio e Rocca, più elegante dal punto di vista teorico, ha caratteristiche di maggior generalità; esso contiene al suo interno, come soluzione semplificata, l'algoritmo di Limb e Murphy.

Sia $u_k(X, Y)$ il valore di luminanza del punto di coordinate x,y nel quadro k; tale punto appartenga ad un oggetto in movimento. Si assume che l'oggetto non cambi la sua luminosità da un quadro all'altro; se questo non è verificato, la stima del movimento può risultare erronea (si tratta comunque di un problema che accomuna praticamente tutti gli algoritmi).

Nel caso di un moto puramente traslatorio, che genera un vettore movimento S avente componenti dx e dy, le differenze tra due quadri successivi valgono:

$$[2] \qquad \begin{aligned} DQ(x,y) &= u_k(x,y) - u_{k-1}(x,y) = \\ &= u_k(x,y) - u_k(x+dx, y+dy) = \\ &= -\frac{\partial u_k(x,y)}{\partial x}dx - \frac{\partial u_k(x,y)}{\partial y}dy - n(x,y) = \\ &= -S^T \nabla u_k(x,y) - n(x,y) \end{aligned}$$

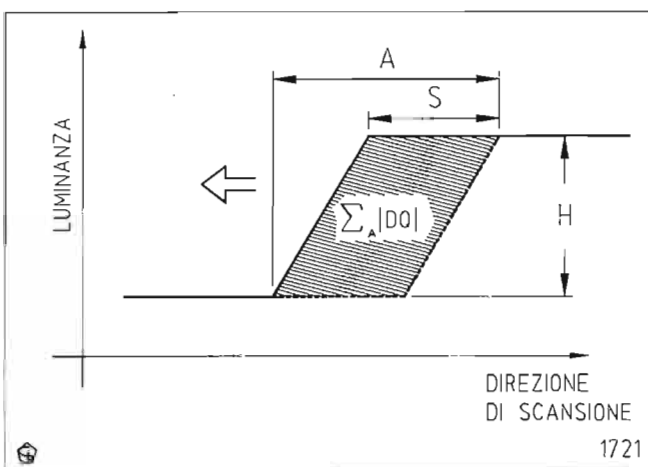
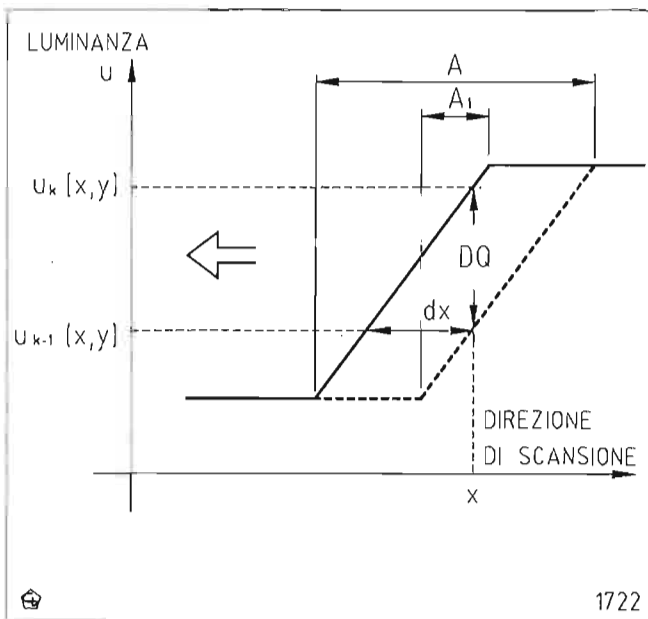



Fig. 2 — Algoritmo di stima del movimento di Cafforio e Rocca. La linea tratteggiata indica la posizione del fronte nel quadro precedente.

dove *n(x,y)* contiene i termini di ordine superiore dello sviluppo di Taylor; nel seguito essi verranno trascurati. Il gradiente ∇u ha per componenti le differenze tra pel vicini della stessa riga video e le differenze tra righe consecutive.

Nella figura 2 è mostrato un caso semplice di spostamento unidimensionale, lungo la sola direzione *x*. In questa situazione, la stima del vettore movimento ottenuta dalla [2] si riduce a

$$\hat{S} = \hat{dx} = - \frac{DQ(x,y)}{\partial u(x,y)/\partial x} \quad [3]$$

L'esattezza di tale risultato può essere immediatamente verificata sulla figura 2; in pratica, si ottiene l'ampiezza *dx* del movimento dividendo la lunghezza del segmento verticale *DQ* per la pendenza del fronte.

Nel caso bidimensionale, se il contorno dell'oggetto che si muove è noto (si veda a proposito il paragrafo 2.1), la [2] può essere valutata su tutti i punti che lo compongono. Impiegando la regressione lineare, è possibile costruire la stima ottimale del vettore movimento *S* come

$$\hat{dx} = - \frac{E\left\{DQ(x,y)\ \frac{\partial u(x,y)}{\partial x}\right\}}{E\left\{\left(\frac{\partial u(x,y)}{\partial x}\right)^2\right\}}$$

$$\hat{dy} = - \frac{E\left\{DQ(x,y)\ \frac{\partial u(x,y)}{\partial x}\right\}}{E\left\{\left(\frac{\partial u(x,y)}{\partial y}\right)^2\right\}} \quad [4]$$

dove le medie sono calcolate sommando sull'intera area *A* occupata dall'oggetto in movimento.

L'algoritmo di Cafforio e Rocca assume che la funzione di luminanza sia lineare nel punto di coordinate *x,y*. Nell'esempio mostrato in figura 2, tale assunzione è valida solo all'interno dell'area A_I; d'altra parte, quest'area diminuisce di dimensioni all'aumentare dello spostamento del fronte.

In definitiva, i due algoritmi visti possono essere usati solo per misurare movimenti *S* molto piccoli (com'era intuibile, d'altronde, per il fatto che adottano un approccio di tipo differenziale). Nel tentativo di superare questo problema, sono state introdotte le tecniche che descriveremo nel seguito.

2.2.1 Tecniche recursive

Per migliorare l'accuratezza della stima del movimento ed allo stesso tempo aumentare il campo dei valori di *S* misurabili, è possibile impiegare algoritmi di tipo recursivo. A partire da una stima iniziale $\hat{S}_i$, si genera una nuova stima migliorata $\hat{S}_{i+1}$ secondo la formula:

$$\hat{S}_{i+1} = \hat{S}_i + T_i \quad [5]$$

dove T_i è il termine di aggiornamento dell'iterazione *i*-esima. Le iterazioni possono avvenire sia su campioni successivi appartenenti alla stessa riga, sia su righe o quadri successivi. A seconda dei casi, si parlerà di algoritmi di stima recursivi con recursione sui campioni orizzontale, verticale o temporale.

Conoscendo $\hat{S}_i$, è possibile utilizzare una funzione delle differenze di quadro traslate *DQT*

$$DQT(x,y,\hat{S}_i) = u_k(x,y) - u_{k-1}(x - \hat{dx}_i, y - \hat{dy}_i) \quad [6]$$

come criterio per il calcolo della stima $\hat{S}_{i+1}$.

Netravali e Robbins (bibl. 6) hanno proposto un algoritmo che tenta di minimizzare recursivamente le differenze al quadrato mediante il metodo del gradiente:

$$\hat{S}_{i+1} = \hat{S}_i - \frac{1}{2} \epsilon \Delta_{\hat{S}_i} [DQT(x,y,\hat{S}_i)]^2 \quad [7]$$

dove $\Delta\hat{S}_i$ rappresenta l'operatore gradiente rispetto a $\hat{S}_i$ ed ϵ è una costante positiva. La figura 3 illustra come

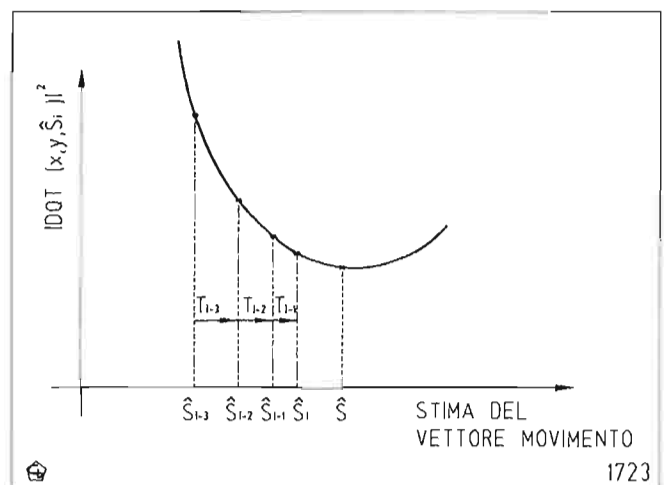



Fig. 3 — Algoritmo di stima del movimento di Netravali e Robbins.

questo algoritmo converga al valore vero del movimento S dopo alcune iterazioni.

La scelta di ϵ è basata su di un compromesso. Valori di ϵ alti portano ad una convergenza rapida ma danno una stima rumorosa, mentre valori più bassi consentono una stima più accurata.

L'algoritmo riassunto nella formula [7] può essere migliorato ricavando il termine di aggiornamento da una media pesata su campioni adiacenti; in questo modo il rumore di quantizzazione viene ad avere un effetto minore. Se A è l'area contenente i campioni considerati, si ottiene:

[8] $$T_i = -\frac{1}{2}\epsilon \Delta_{\hat{S}_i} \sum_{j \in A} W_j [DQT(x,y,\hat{S}_i)]^2$$

dove $W_j \geq 0$ e $\Sigma W_j = 1$.

In seguito al lavoro di Netravali e Robbins, sono stati sviluppati molti altri algoritmi recursivi; la differenza principale consiste nel sostituire la costante ϵ con una variabile, in modo da permettere un migliore adattamento alle caratteristiche locali dell'immagine. Il risultato di tale modifica è una velocità di convergenza più elevata, ovvero una migliore precisione di stima quando il numero di iterazioni sia fissato a priori.

Nonostante gli algoritmi recursivi funzionino bene per movimenti di pochi pel o addirittura di frazioni di pel, essi possono incorrere in situazioni di instabilità quando l'immagine è molto rumorosa oppure il movimento è eccessivamente veloce. Inoltre, volendo passare ad una fase realizzativa, occorre ricorrere a tecniche di interpolazione e stima semplificate, che non garantiscono sempre risultati ottimali. Pertanto, il loro uso all'interno di un codificatore video a riduzione di ridondanza è fortemente penalizzante, e non è stato considerato al *Centro Ricerche* RAI.

2.3 TECNICHE BASATE SULLA TRASFORMATA DI FOURIER

La procedura di base, per questi algoritmi, consiste nel determinare la superficie di correlazione tra due immagini (due quadri, o semiquadri, successivi); tale superficie presenterà un picco in corrispondenza alle coordinate che danno lo spostamento di un'immagine rispetto all'altra. Naturalmente è possibile che diverse parti dell'immagine, per esempio gli oggetti che sono presenti nella scena, si muovano in maniere differenti; in quest'ultimo caso, che d'altra parte è il più frequente, la superficie di correlazione avrà un certo numero di picchi, ognuno dei quali è associato al movimento di una determinata area d'immagine.

La tecnica impiegata per valutare la funzione di correlazione si basa sull'uso della trasformata di Fourier. Come primo passo, alle due immagini viene applicata una trasformazione bidimensionale; i due spettri ottenuti vengono moltiplicati tra di loro, quindi la funzione risultante viene antitrasformata sempre secondo Fourier. Il risultato è la superficie di correlazione desiderata.

I vantaggi ottenibili, rispetto al calcolo diretto della funzione di correlazione, stanno nel minore impegno di calcolo e nella possibilità, molto importante, di applicare un filtraggio sulla superficie di correlazione. In particolare, è molto usata la tecnica di normalizzare l'ampiezza dello spettro risultante dal prodotto, in modo da preservare le sole componenti di fase: se G_1 e G_2 sono le trasformate discrete di Fourier (DFT = Discrete Fourier Transform) delle due immagini in considerazione, la superficie di correlazione $Y(x,y)$ si ottiene come antitrasformata della funzione complessa Z, data da:

[9] $$Z = \frac{G_1 G_2^*}{|G_1 G_2^*|}$$

Mediante questo algoritmo, noto anche come **correlazione di fase**, si ottiene una stima del movimento molto più precisa; i picchi aumentano di ampiezza rispetto al resto della superficie, e risultano pertanto più facilmente individuabili. Ciò è dovuto al fatto che tutta l'informazione di posizione è contenuta nella fase dello spettro, mentre eventuali variazioni di luminosità degli oggetti, che potrebbero degradare la misura, vengono annullate dalla normalizzazione del modulo.

Le prestazioni di tale tecnica sono molto buone; possono essere stimati movimenti di molte decine di pel per quadro, con precisioni che possono arrivare a meno di un decimo di pel (ciò se si effettua un'interpolazione della superficie di correlazione). Il problema principale dell'algoritmo, ossia la necessità di separare i movimenti delle diverse parti dell'immagine, può essere risolto mediante vari accorgimenti; su questi ultimi non ci soffermiamo, ma rimandiamo alla bibliografia 7.

Nel complesso, si tratta di una tecnica relativamente nuova, sulla quale sono necessari studi ulteriori; essa appare comunque di notevole interesse, e potrebbe in futuro trovare utili applicazioni nei sistemi di codifica a basso bit-rate (minore di 1 bit/pel).

2.4 BLOCK MATCHING

Gli algoritmi recursivi per la stima dei vettori movimento si basano sul calcolo del gradiente; in alternativa a ciò, lo spostamento subito da una parte d'immagine può essere valutato dalla superficie di correlazione tra due quadri successivi. È possibile definire una funzione bidimensionale di mutua correlazione C come:

[10] $$C(S) = \frac{R_{u_k u_{k-1}}(S)}{\sqrt{R_{u_k u_k}(0) R_{u_{k-1} u_{k-1}}(0)}}$$

(ricordiamo che S rappresenta il vettore movimento, formato dalle due componenti dx e dy). La stima $\hat{S}$ dello spostamento è data dalla posizione del picco di correlazione.

Per trovare lo spostamento subito dal pel di coordinate x,y occorre estrarre dal quadro k-esimo un blocco di $M \cdot N$ pel centrato sul punto x,y, quindi effettuarne la correlazione con i pel appartenenti ad una certa area di ricerca AR nel quadro precedente $(k-1)$. Poiché le dimensioni dell'area di ricerca superano quelle del blocchetto di un'entità corrispondente al massimo spostamento previsto tra quadri successivi, come indicato dalla figura 4, e dato che il blocchetto deve essere piuttosto grande (almeno 8·8 pel) al fine di ridurre la varianza della stima, la quantità di calcoli richiesta per la determinazione del movimento risulta eccessivamente elevata.

Una soluzione molto efficace per ridurre la complessità dell'algoritmo consiste nel suddividere l'immagine in blocchi, considerando che tutti i pel appartenenti allo stesso blocco abbiano vettori movimento coincidenti. In que-

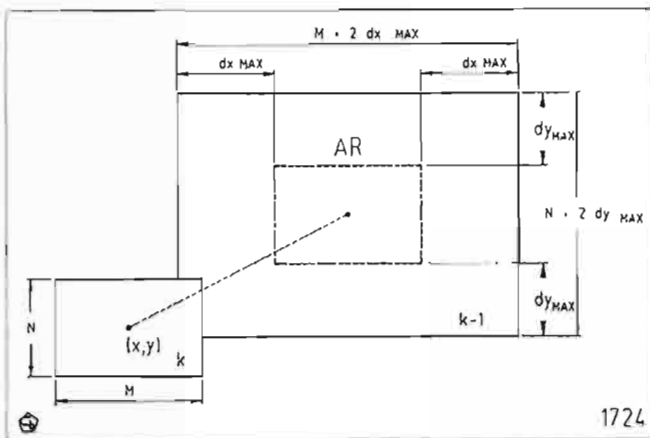



Fig. 4 — Posizione del blocco di M·N pel, appartenente al quadro k, all'interno dell'area di ricerca AR nel quadro $k-1$.

sto modo, occorre calcolare un solo vettore per ogni blocco. La tecnica di correlazione così semplificata prende il nome di *block matching*.

Come si può notare, una tecnica di questo tipo è particolarmente adatta ad essere applicata nell'algoritmo DCT ibrido; in effetti, l'immagine risulta già di per sé suddivisa nei blocchi di 8·8 pel sui quali viene calcolata la trasformazione. Per ciascun blocco, la predizione può essere ottenuta dal semi-quadro o dal quadro precedente, applicando una traslazione di entità pari al vettore movimento calcolato per il blocchetto.

In particolare, gli studi e le simulazioni eseguite hanno fatto preferire, in prima istanza, la compensazione movimento inter-frame; essa, infatti, non risente del disturbo dovuto all'interallacciamento, e dà comunque buoni risultati, nonostante la distanza temporale tra due quadri successivi sia doppia rispetto a quella che separa due field.

Una semplificazione ulteriore della ricerca dei vettori movimento, d'altra parte, è ottenuta impiegando come criterio di correlazione una funzione differente da quella definita nella [10]; essa risulta infatti troppo onerosa dal punto di vista dei calcoli. Si è scelto di utilizzare lo stesso criterio già impiegato nella scelta «a priori» del modo di trasmissione, ossia l'errore quadratico medio

$$[11] \quad \sigma^2(i,j) = \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} [u_k(m,n) - u_{k-1}(m+i,\ n+j)]^2$$

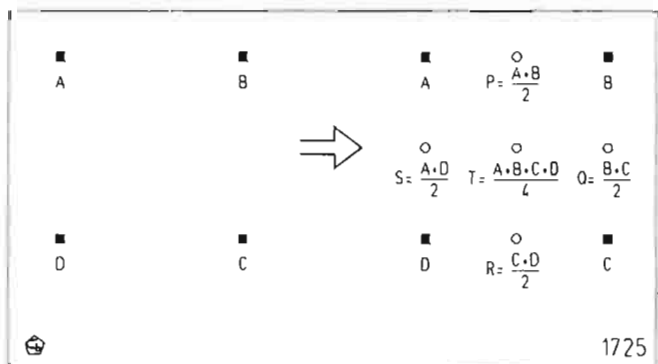



Fig. 5 — Costruzione delle interpolazioni a 1/2 pel per la compensazione movimento.

Le dimensioni del blocco sono state fissate pari a 16·8 pel (tali valori corrispondono in effetti ad un *quadriblocco*, ovvero un'area di immagine la cui forma sullo schermo risulta approssimativamente quadrata). Lo spostamento massimo considerato è di ±15,5 pel nella direzione orizzontale e ±7,5 pel nella direzione verticale. Può apparire strano l'uso di coordinate non intere; in realtà, per migliorare la precisione della stima ed al tempo stesso riuscire a seguire movimenti molto lenti, prima del *matching* si effettua un'interpolazione lineare bidimensionale, in modo da portare la risoluzione a 1/2 pel.

La tecnica usata è mostrata in figura 5. Essa risulta particolarmente semplice, pertanto è adatta ad un'immediata realizzazione hardware (nell'ambito del progetto EU 256).

È da notare, tuttavia, che un'interpolazione lineare di questo tipo non dà luogo alla ricostruzione ottimale dei campioni mancanti; in effetti, potrebbe essere conveniente impiegare filtri di interpolazione più complessi. Ciò è attualmente oggetto di studio, insieme alla valutazione degli eventuali vantaggi nell'uso collaterale dell'informazione inter-field per la stima del movimento.

Negli algoritmi di *block matching*, si sono considerati tradizionalmente tre sistemi di ricerca del picco di correlazione:

esaustivo. Come si può comprendere dalla denominazione, questo metodo si basa sulla «forza bruta»; per ognuno dei blocchi in cui è stata divisa l'immagine, vengono provate tutte le possibili posizioni all'interno dell'area di ricerca. Al termine, si sceglie come stima del vettore movimento il valore corrispondente alla posizione di massima correlazione (nel caso in cui venga usato l'errore quadratico medio, la posizione ottimale sarà invece quella corrispondente al minimo). Il difetto principale di questo approccio consiste ovviamente nella grande quantità di calcoli richiesti; d'altra parte, esso garantisce l'individuazione corretta del picco e quindi del vettore movimento;

2-D logaritmico. Il funzionamento di questo algoritmo, proposto da J. R. Jain e A. K. Jain, si può comprendere osservando la figura 6. Ad ogni passo, si calcola l'errore quadratico medio in cinque posizioni; quindi si procede lungo la direzione associata al punto che ha dato il risultato migliore. Nel caso in cui il minimo sia stato trovato sul punto centrale, oppure se si sono raggiunti i confini dell'area di ricerca (nella figura 6 ciò avviene al passo 4) si procede riducendo la distanza tra i punti (bibl. 8). Volendo determinare uno spostamento massimo di 6 pel, come nell'esempio mostrato in figura 6, sono richiesti nel caso peggiore 7 passi, corrispondenti al calcolo dell'errore su 21 punti. L'algoritmo esaustivo, nella stessa situazione, dovrebbe effettuare il calcolo su 169 punti;

three-step. Pubblicato da Koga *et al.*, l'algoritmo *three-step* (figura 7) prevede un test iniziale su otto punti, collocati simmetricamente attorno al centro $x=i$, $y=j$ e spaziati in modo piuttosto largo. Al secondo passo, la distanza tra i punti viene ridotta, mentre il centro viene posto sulla posizione migliore determinata al primo tentativo. Si procede iterativamente in questo modo, fino a quando la spaziatura non può essere ulteriormente ridotta (bibl. 8). Nel caso particolare di spostamento massimo pari a ±6 pel, l'algoritmo termina esattamente in tre passi, da qui il nome con cui viene indicato. Rispetto al metodo logaritmi-

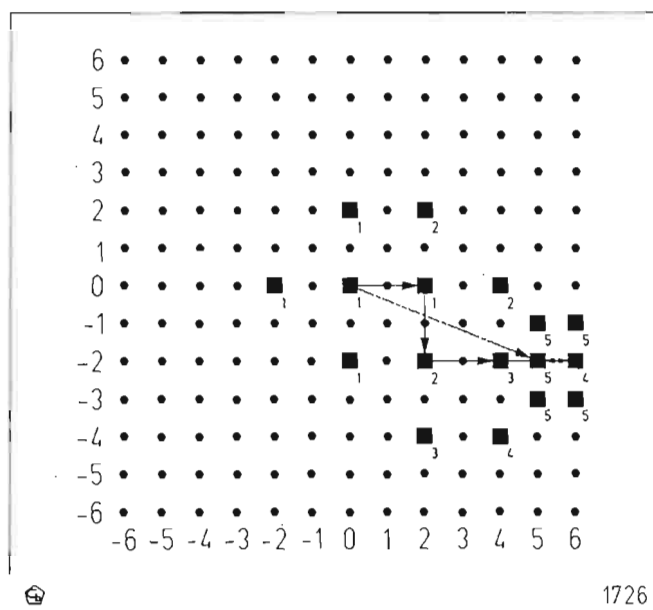



Fig. 6 — Metodo 2D-logaritmico. Sono mostrati i punti di ricerca nel quadro $k-1$, relativi al pel di coordinate (i,j) nel quadro k.

co, la situazione peggiore comporta la prova di un numero più alto di punti (25); tuttavia, nel caso di elaborazione parallela, il fattore più importante è dato dal numero di passi successivi, che nel *three-step* è mediamente inferiore.

Gli algoritmi 2-D logaritmico e *three-step* operano una riduzione consistente della complessità computazionale, in quanto determinano il vettore movimento attraverso il calcolo di un numero molto limitato di punti della superficie di correlazione. Tuttavia, come altre tecniche analoghe di ricerca dei punti di massimo o di minimo di una funzione, esse sono basate sull'ipotesi fondamentale di una notevole regolarità della superficie analizzata.

Se tale ipotesi non è verificata, il successo della ricerca non è garantito. Ad esempio, nel caso dell'algoritmo *three-step*, la funzione può assumere nell'intorno del punto centrale un andamento irregolare, che porta alla scelta di una direzione di partenza sbagliata. Una volta decisa tale direzione, d'altra parte, non è più possibile effettuare alcuna correzione; la ricerca prosegue in una zona che non contiene affatto il punto di minimo. Di conseguenza si ottiene una stima del vettore movimento completamente errata.

La possibilità di sbagliare direzione, in effetti, sussiste ad ogni passo di iterazione; l'algoritmo converge sicuramente solo se la superficie di correlazione ha una forma molto regolare.

Contrariamente a quanto si potrebbe credere, questa regolarità è in realtà poco probabile. Si osserva infatti che l'andamento tipico della funzione di errore quadratico medio, nel caso di movimento puramente traslatorio, ha l'aspetto mostrato in figura 8: la funzione presenta un avvallamento molto pronunciato in corrispondenza dello spostamento subito dall'area di immagine considerata, ma al di fuori di tale avvallamento le ondulazioni sono del tutto casuali. Esistono inoltre numerosi punti di minimo locale, che non coincidono con il minimo assoluto. In altre parole, non è possibile avvicinarsi al minimo seguendo l'andamento globale della funzione.

Per i motivi succitati, la tecnica di compensazione movimento adottata nel codificatore RAI-Telettra prevede la ricerca esaustiva; prove sperimentali e considerazioni teoriche hanno dimostrato che i vantaggi ottenibili in termini di velocità dalle altre tecniche non sono sufficienti a giustificare la perdita considerevole di efficienza che comportano. Dal punto di vista della realizzazione hardware, il problema della complessità di calcolo può essere risolto facendo ricorso ad un'architettura con un elevato grado di parallelismo; purtroppo, però, in fase di simulazione la quantità di calcoli richiesta rende i tempi di attesa eccessivamente lunghi. In effetti, l'area di ricerca prevista, pari a $\pm 15,5$ pel in orizzontale e $\pm 7,5$ pel in verticale, implica la valutazione dell'errore quadratico medio in 1953 posizioni diverse; questo per ognuno dei 1620 quadriblocchi che compongono un semiquadro 4:2:2 (la situazione, ovviamente, è ancora peggiore nel caso dell'alta definizione).

Nonostante la valutazione del vettore movimento sia fatta solo sulla base della luminanza (soluzione che, pur non essendo ottimale, si è dimostrata molto vantaggiosa in termini di rapporto efficienza/complessità), la simulazione dell'algoritmo DCT ibrido con compensazione movimento richiede, quando si usi la tecnica esaustiva, tempi dell'ordine di un'ora per ogni quadro della sequenza.

Per risolvere questo problema, gli autori hanno sperimentato con successo un metodo basato sulla predizione dei vettori movimento, in grado di ridurre di un fattore 6-8 i tempi di simulazione; di tale metodo diamo la descrizione completa in appendice.

3. Valutazioni quantitative

Per mezzo delle apparecchiature disponibili al *Centro Ricerche* RAI (bibl. 10) è stato eseguito un grande nume-

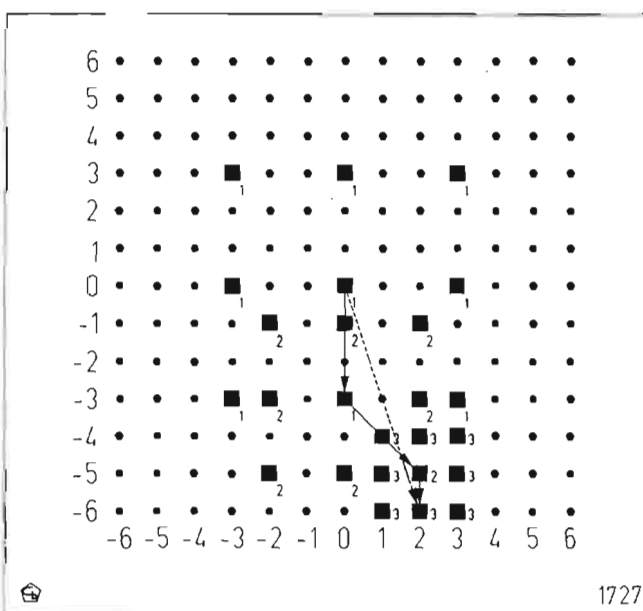
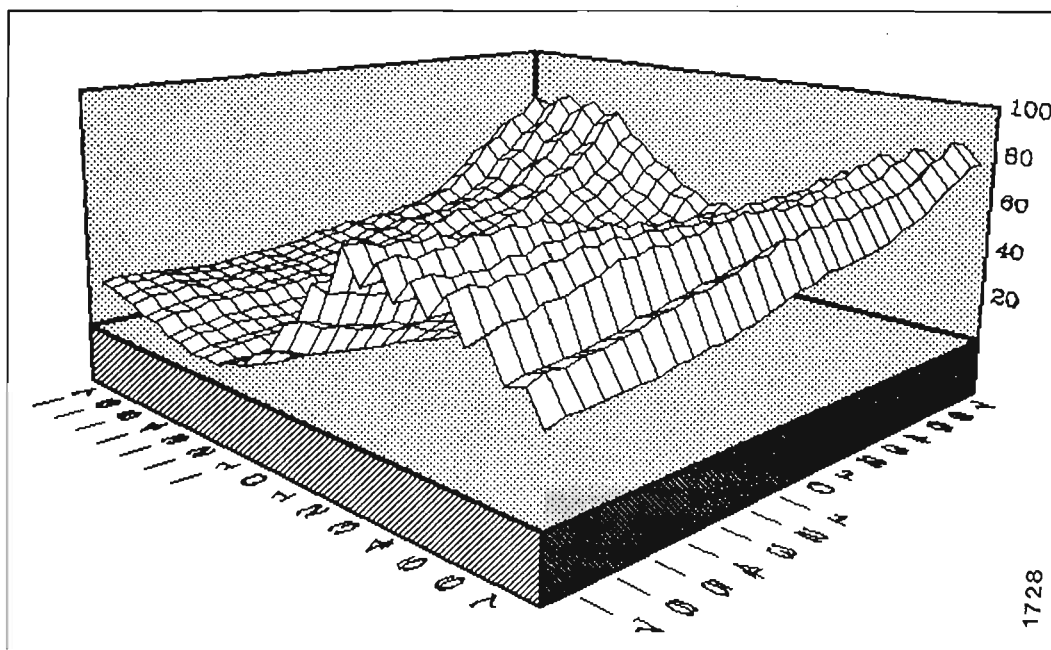



Fig. 7 — Metodo *three-step*. Sono mostrati i punti di ricerca nel quadro $k-1$, relativi al pel di coordinate (i,j) nel quadro k.

Fig. 8 — Andamento tipico dell'errore quadratico medio calcolato tra blocchi corrispondenti di immagine appartenenti a due quadri successivi, al variare della posizione di uno dei blocchi. La figura mostra una variazione di posizione massima pari a ±7,5 pel, nelle direzioni orizzontale e verticale.

ro di prove di simulazione, in modo da verificare le effettive prestazioni della tecnica di compensazione movimento, riportiamo nel presente lavoro i risultati più significativi.

La Tabella 1 mostra i guadagni di bit-rate ottenibili con l'uso della compensazione movimento, qualora si operi a fattore di trasmissione costante (in pratica, mantenendo costante la qualità dell'immagine ricostruita). Come si può vedere, si tratta di guadagni, piuttosto consistenti, che possono anche superare il 30%; sulla sequenza «Table Tennis» la riduzione non è così elevata, ma occorre notare che essa risulta già di per sé poco critica, anche senza compensazione.

Nella Tabella 1 sono evidenziate due possibilità di ricerca del vettore movimento, la prima prevede una risoluzione pari ad 1 pel (il block matching viene effettuato direttamente sui punti dell'immagine), mentre la seconda avviene mediante una preventiva interpolazione lineare, che porta la risoluzione a 1/2 pel (vedere il paragrafo 2.4). Si osserva che l'aumento di risoluzione porta ad un considerevole miglioramento delle prestazioni, sufficiente a giustificare la maggiore complessità di realizzazione.

La Tabella 1 riporta altresì le distribuzioni statistiche dei modi di trasmissione, distinte per luminanza e crominanza; come era lecito aspettarsi, l'applicazione della compensazione movimento inter-frame aumenta enormemente le percentuali di scelta di quest'ultima modalità di codifica.

Sulla crominanza l'aumento risulta più contenuto; ciò è dovuto principalmente al fatto che la valutazione del vettore movimento avviene sulla base della sola informazione di luminanza. Tuttavia non si tratta di un problema di particolare conto, in quanto il bit-rate necessario per

TABELLA 1.
GUADAGNI DI BIT-RATE OTTENIBILI CON LA COMPENSAZIONE MOVIMENTO

Sequenza	Comp. movimento	Bit-rate (Mbit/s)	Guadagno (%)	Modi Y (%) Intra field	Modi Y (%) Inter field	Modi Y (%) Inter frame	Modi C (%) Intra field	Modi C (%) Inter field	Modi C (%) Inter frame
Renata	No	21,59	0	29,7	43,0	27,3	35,1	45,9	19,0
	1 pel	16,78	22,23	9,8	15,2	75,0	22,2	24,1	53,7
	1/2 pel	14,78	31,52	5,1	8,0	86,9	22,6	23,4	54,0
Mobile & Calendar	No	30,12	0	8,9	72,8	18,3	9,8	70,7	19,5
	1 pel	22,04	26,83	5,0	13,7	81,3	5,8	20,0	74,2
	1/2 pel	19,65	34,77	2,9	9,2	87,9	5,6	19,3	75,1
Table Tennis	No	16,04	0	25,3	33,1	41,6	34,2	36,1	29,7
	1 pel	14,32	10,71	16,1	20,4	63,5	27,8	27,0	45,2
	1/2 pel	13,29	17,13	12,5	13,9	73,6	26,8	25,5	47,7

NOTA: nei due casi la risoluzione della stima è pari rispettivamente a 1 pel e 1/2 pel. Le simulazioni sono state effettuate a fattore di trasmissione costante e quindi a qualità costante.

la codifica delle componenti di crominanza è normalmente una porzione molto esigua del bit-rate complessivo.

TABELLA 2.
MIGLIORAMENTI QUALITATIVI OTTENIBILI CON LA COMPENSAZIONE MOVIMENTO

Sequenza	Bit-rate (Mbit/s)	con comp. movimento S/N (dB)	senza comp. movimento S/N (dB)
Flower Garden	9,5	31,35	26,11
	12	33,04	28,11
	15,5	34,87	30,24
Renata	9,5	32,66	29,15
	12	33,95	30,90
	15,5	35,37	32,72
Table Tennis	9,5	34,39	32,68
	12	35,22	33,96
	15,5	36,21	35,24

NOTA: le simulazioni sono state effettuate a bit-rate costante.

La Tabella 2 presenta una situazione complementare: operando a bit-rate costante, si cerca di valutare il miglioramento qualitativo conseguente all'uso della compensazione movimento.

Anche questa volta, le aspettative trovano conferma in un aumento consistente del rapporto segnale/rumore, che può arrivare a più di 5 dB.

TABELLA 3.
GUADAGNI DI BIT-RATE E RAPPORTI S/N DELL'IMMAGINE RICOSTRUITA CON L'USO DELLA COMPENSAZIONE MOVIMENTO

Sequenza	Comp. movimento	Bit-rate (Mbit/s)	Guadagno (%)	S/N (dB)
Mobile & Calendar	No	30,12	0	33,54
	Esaustiva	19,65	34,77	33,81
	Ridotta	19,78	34,33	33,80
Renata	No	21,59	0	35,04
	Esaustiva	14,78	31,52	35,47
	Ridotta	15,26	29,31	35,44
Table Tennis	No	16,04	0	35,53
	Esaustiva	13,29	17,13	35,73
	Ridotta	14,13	11,88	35,69
Renata & Butterflies	No	22,94	0	34,13
	Esaustiva	19,68	14,19	34,48
	Ridotta	20,04	12,61	34,44
Diva with Noise	No	75,02	0	27,34
	Esaustiva	73,44	2,09	27,36
	Ridotta	74,31	0,93	27,36

NOTA: l'immagine è ricostruita con l'uso della compensazione movimento nei due casi: di ricerca esaustiva o ridotta del vettore movimento. Le simulazioni sono state effettuate a fattore di trasmissione costante.

La Tabella 3, analogamente alla Tabella 1, presenta i risultati di simulazioni eseguite a qualità costante; questa volta, però, l'accento è posto principalmente sulla validità del metodo di simulazione fondato sulla predizione inter-field dei vettori movimento, che opera la ricerca di questi ultimi nell'ambito di una finestra ridotta. Come si vede, le prestazioni sono del tutto confrontabili con quelle della ricerca esaustiva, tranne nel caso della sequenza «Table Tennis»; il motivo per cui questa sequenza causa dei problemi, ed il modo di risolverli, verranno trattati in appendice.

Un altro aspetto interessante dei dati in Tabella 3 riguarda la sequenza «Diva with Noise»; come si è già avuto occasione di dire (bibl. 10) si tratta di un'immagine in cui la maggior parte dello schermo è occupata da rumore bianco a pieno livello, sia sulla luminanza che sulla crominanza. Ovviamente, nella parte contenente rumore non ha alcun senso cercare di determinare dei vettori movimento; per questo motivo i guadagni ottenibili sono estremamente ridotti, e riguardano solo la zona predicibile dell'immagine. Si tratta, comunque, di un caso costruito *ad hoc* per «mettere in crisi» l'algoritmo di codifica.

4. Conclusioni

In passato, l'attenzione del *Centro Ricerche* RAI riguardo alla codifica del segnale video digitale è stata volta principalmente alle esigenze di collegamento tra studi televisivi; in collaborazione con la Telettra, è stato sviluppato un codec in grado di garantire sul segnale la *qualità contributo*, ovvero una precisione di ricostruzione sufficiente per eseguire alcune operazioni di post-processing particolarmente critiche, come ad esempio il *Chroma-key*.

In queste condizioni operative, l'uso della compensazione movimento va inteso non tanto come mezzo di riduzione del bit-rate, quanto come una tecnica che permette di mantenere ottimi standard qualitativi anche su sequenze molto critiche (immagini con forte contenuto di dettaglio spaziale e variabilità temporale). In effetti, il flusso di informazione risulta stabilito dalla gerarchia di multiplazione che si intende utilizzare.

Nella situazione attuale, tuttavia, sempre maggiore attenzione è attribuita, da parte dei gruppi internazionali, alla *distribuzione* del segnale all'utente finale ed alla sua registrazione su cassette mediante video-registratore domestico. In questo caso, le esigenze di qualità sono meno stringenti mentre risulta importante ridurre il più possibile il flusso di dati binari.

Inoltre, assumono maggiore importanza le prestazioni soggettive dell'algoritmo (visibilità del rumore di codifica, eventuale presenza di *artefact*), rispetto a quelle oggettive (rapporto S/N dell'immagine decodificata).

La compensazione movimento, in questo campo, assume un'importanza notevole per il guadagno di bit-rate che consente di ottenere; essa andrà comunque usata in congiunzione ad altre tecniche operative, in modo da raggiungere traguardi che solo pochi anni fa sarebbero apparsi eccessivamente ambiziosi (codifica del segnale 4:2:2 con un bit-rate complessivo di 5 Mbit/s).

APPENDICE

Predizione dei vettori movimento

Il concetto che sta alla base di questo algoritmo nasce dall'esame delle statistiche effettuate dagli autori sulla distribuzione dei vettori movimento; per diverse sequenze di prova, rappresentative di una normale trasmissione televisiva, sono stati valutati gli istogrammi dei vettori movimento e quelli delle corrispondenti differenze inter-field (figura 1A). Queste ultime sono definite nel modo seguente: se S_k e S_{k-1} sono i vettori movimento relativi rispettivamente ai semi-quadri k e k_{-1}, la differenza inter-field (che si può altresì interpretare come errore di predizione del vettore al semi-quadro k) vale

$$e_k = S_k - S_{k-1} \tag{12}$$

Dalle statistiche è risultato che la distribuzione degli errori di predizione e_k presenta un picco molto pronunciato in corrispondenza dello zero; inoltre essa decresce con notevole rapidità all'aumentare del valore di e_k. Ciò può essere interpretato in modo molto semplice, osservando che gli oggetti in una scena televisiva, pur potendo presentare talvolta movimenti piuttosto veloci, non subiscono generalmente accelerazioni molto forti.

Grazie a questa proprietà, la ricerca dei vettori movimento S_k relativi ad un semiquadro può essere resa notevolmente più efficiente, se si adotta l'accorgimento di limitare la *Search Area* ad un intorno conveniente della posizione data dal vettore movimento S_{k-1}, calcolato nel semi-quadro precedente.

Questa tecnica permette una riduzione consistente della «finestra di ricerca»; le simulazioni effettuate hanno mo-

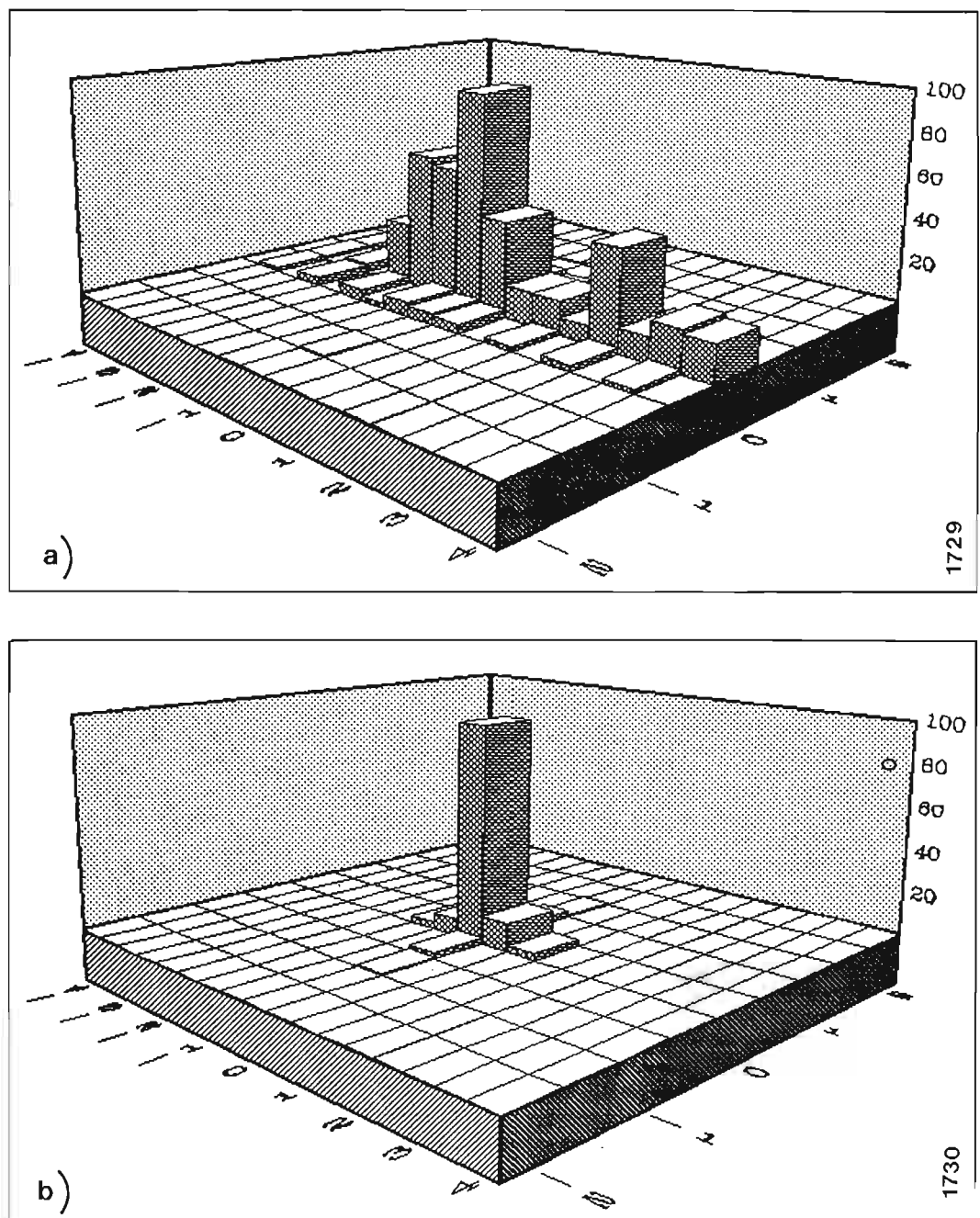



Fig. 1A — *a)* istogramma dei vettori movimento per la sequenza Renata; *b)* istogramma delle differenze inter-field dei vettori movimento per la stessa sequenza. In entrambi i casi la finestra considerata ha estensione pari a ±4 pel in orizzontale e ± 2 pel in verticale.

strato come sia sufficiente una finestra di 4 pel in orizzontale e 2 pel in verticale, ossia un numero di posizioni pari ad 1/16 di quelle richieste dal metodo esaustivo.

Le prestazioni ottenibili sono molto buone; su sequenze di durata pari a 100 quadri (4 secondi) la perdita di efficienza, misurata in termini di aumento del bit-rate complessivo necessario alla codifica, è in media di 2%, come si può vedere dalla Tabella 3.

Il motivo per cui la tecnica predittiva dà risultati migliori, rispetto ad algoritmi come il three-step, sta nel fatto che, all'interno della finestra ridotta, la ricerca avviene in modo esaustivo. Se il picco di correlazione non si sposta eccessivamente da un semi-quadro al successivo (in altre parole, se rimane all'interno della finestra) esso viene **sicuramente** individuato; lo stesso non si può dire degli altri algoritmi, che scelgono arbitrariamente dei punti di prova senza tener conto delle informazioni già disponibili.

Da quanto detto, risulta evidente che in casi molto particolari la tecnica può fallire: ad esempio, questo succede quando vi sia un brusco cambiamento di scena, per cui i valori di predizione non hanno significato, oppure in presenza di accelerazioni elevate. In effetti, quando ciò avviene, per ritornare ad un funzionamento corretto può essere richiesto un tempo corrispondente a molti field; è il caso della sequenza «Table Tennis» nella Tabella 3, in cui la presenza di un cambiamento di scena rende più critica l'applicazione dell'algoritmo (il bit-rate aumenta circa del 6% rispetto alla ricerca esaustiva).

TABELLA 4.
PRESTAZIONI DELL'ALGORITMO DI RICERCA PREDITTIVA DEI VETTORI MOVIMENTO AL VARIARE DEL PERIODO DI *REFRESH*

Sequenza	Periodo di refresh (semi-quadri)	Guadagno (%)	S/N (dB)
Renata	10	30,49	35,45
	20	30,16	35,45
	30	29,84	35,45
Table Tennis	10	16,12	35,72
	20	15,51	35,71
	30	14,87	35,71

NOTA: distanza in semi-quadri tra due applicazioni consecutive della ricerca esaustiva.

Per eliminare questo problema, sono possibili diverse soluzioni; una strategia molto semplice, ma efficace, consiste nell'eseguire una sorta di *refresh*, allargando periodicamente la finestra in modo da individuare con certezza la posizione del picco.

Il periodo di refresh deve essere scelto sufficientemente basso, in modo da evitare che l'aggancio della finestra con il vettore movimento possa venire a mancare per troppi semi-quadri consecutivi; d'altra parte, per non aumentare eccessivamente i tempi di simulazione, sarebbe conveniente allungare il più possibile la distanza tra due refresh successivi.

Le simulazioni eseguite (vedere la Tab. 4) hanno mostrato come un periodo di refresh pari a 20 semi-quadri offra un compromesso assai soddisfacente tra velocità e prestazioni dell'algoritmo.

(3953)

BIBLIOGRAFIA

1. - MAY F., WOLF W.: *Picture coding with motion analysis for low bit rate transmission*, ICC 82, Proc., pagg. 2G.7.1-2G.7.5, Giugno 1982.
2. - LABIT C.,BENVENISTE A.: *Motion estimation in a sequence of television pictures*, «Image Sequence Processing and Dynamic Scene Analysis», T. S. Huang, Ed. Springer-Verlag, 1983.
3. - KRETZ F.: *Edges in visual scenes and sequences: applications to filtering, sampling and adaptive DPCM coding*, «Image Sequence Processing and Dynamic Scene Analysis», T. S. Huang, Ed. Springer-Verlag, 1983.
4. - LIMB J.O., MURPHY J. A.: *Measuring the speed of moving objects from television signals*, IEEE Trans. Commun., vol. COM-23, n. 4, Aprile 1975.
5. - CAFFORIO C., ROCCA F.: *Methods for measuring small displacements of television images*, IEEE Trans. Inform. Theory, vol. IT-22, n. 5, Sett. 1976
6. - NETRAVALI A. N., ROBBINS J. D.: *Motion compensated television coding-Part I*, «Bell System Technology Journal», vol. 58, Marzo 1979.
7. - THOMAS G. A., HONS B. A.: *Television motion measurement for DATV and other applications*, BBC Research Department Report n. BBC RD 1987/11.
8. - MUSMANN H. G., PIRSCH P., GRALLERT H. J.: *Advances in picture coding*, Proc. of the IEEE, vol. 73, n. 4, Aprile 1985.
9. - DEL PERO R.: *Codifica del segnale televisivo numerico: metodi di riduzione della ridondanza*, «Elettronica e Telecomunicazioni», n. 1, 1989.
10. - BARBERO M., STROPPIANA M.: *Codifica del segnale televisivo numerico: uso della trasformata coseno discreta*, «Elettronica e Telecomunicazioni», n. 1, 1989.
11. - BARBERO M., CUCCHI S.: *Codifica del segnale televisivo numerico: architettura di un co-decodificatore HDTV utilizzante la DCT*, «Elettronica e Telecomunicazioni», n. 1, 1990.
12. - BELLORA R., DIMINO G., MURATORI M.: *Hybrid DCT: comparison of the statistics of DCT coefficients and processing modes with and without motion compensation*, in «Signal processing of HDTV, 2», Elsevier Science Publischers (Proc. of the 3rd Int. Workshop on HDTV, Torino I, Agosto 1989).
13. - BARBERO M., CUCCHI S., DEL PERO R., DIMINO G., OCCHIENA M., MURATORI M., STROPPIANA M.: *ITALIA '90: Codifica del segnale televisivo numerico, «Elettronica e Telecomunicazioni»*, n. 3, 1990.
14. - BARBERO M., DEL PERO R., RAVERA S., ROSSATO L.: *The use of bit-rate reduction techniques for the transmission of video signals, DSP90* - 2nd Int. Workshop on Digital Signal Processing Techniques Applied to Space Communications, Torino I, Sett. 1990.

CODIFICA DEL SEGNALE TELEVISIVO NUMERICO: QUANTIZZAZIONE E CODIFICA VETTORIALE DEI COEFFICIENTI DCT

L. BOCH, M. STROPPIANA*

SOMMARIO — I metodi basati sulla Trasformata Coseno Discreta per la riduzione della ridondanza del segnale televisivo a definizione convenzionale e del segnale HDTV sono stati oggetto di studio presso il Centro Ricerche della RAI. Gli algoritmi sviluppati utilizzano la quantizzazione scalare dei coefficienti DCT. I coefficienti quantizzati vengono codificati, sempre scalarmente, mediante codici a lunghezza variabile (VLC). Al fine di aumentare l'efficienza delle operazioni di quantizzazione e codifica si sono studiate le tecniche vettoriali e confrontate con le rispettive scalari. Nella prima parte dell'articolo si descrivono tali tecniche vettoriali, mentre nella seconda parte si riportano i risultati ottenuti con l'applicazione di tali metodi all'algoritmo di DCT Ibrida.

*SUMMARY — **Digital TV signal coding: vector quantization and coding of the DCT coefficients:** The methods for the bit-rate reduction of conventional and high definition video signals based on the Discrete Cosine Transform (DCT) have been studied at the RAI Research Centre. The developed algorithm applies scalar quantisation to the DCT coefficients. The quantised coefficients are entropy encoded by means of a Variable Length Code (VLC). To increase the performance of the coding system, vector quantisation and vector coding have been studied and compared with the previously adopted scalar approaches. Firstly the vector quantisation and coding techniques are described and then the results concerning the vector quantisation and coding are reported. A comparison between the scalar and vector operations is also depicted.*

1. Premessa

In precedenti articoli (bibl. 1, 2, 3 e 4) apparsi su questa rivista si sono descritti gli studi relativi all'algoritmo di codifica, mediante DCT (Discrete Cosine Transform), del segnale televisivo numerico. Tali studi svolti nell'ambito del progetto europeo di ricerca EUREKA EU-256[1] portarono alla definizione di una famiglia di co-decodificatori (bibl. 5) per la trasmissione numerica di segnali TV e HDTV con qualità contributo. Alcuni prototipi di co-decodificatori TV e HDTV sono già stati realizzati in hardware e dimostrati ad organismi internazionali, per esempio al CMTT/2 che ha il compito di definire un sistema di codifica per la trasmissione del segnale televisivo conforme alla Raccomandazione 601 del CCIR. I co-decodificatori HDTV, non ancora utilizzanti la compensazione del movimento, sono stati utilizzati per la trasmissione punto-multipunto del segnale HDTV codificato a circa 70 Mbit/s durante i mondiali di calcio «Italia '90» (bibl. 6). L'algoritmo sviluppato di DCT Ibrida, DCT spaziale bi-dimensionale e predizione DPCM (Differential Pulse Code Modulation) temporale, presenta un'elevata efficienza di codifica ed una grande robustezza contro gli errori residui di canale, errori presenti dopo la correzione operata dal FEC (Forward Error Correction). I coefficienti DCT sono quantizzati scalarmente e poi codificati, sempre scalarmente, mediante codici a lunghezza variabile (VLC, Variable Length Code). Prima della codifica a lunghezza variabile dei simboli si ha una codifica delle sequenze di zeri consecutivi lungo il percorso di scansione del blocco 8·8 (bibl. 4).

Al fine di aumentare l'efficienza delle operazioni di quantizzazione e codifica si sono esaminate le tecniche vettoriali. Nella prima parte dell'articolo viene descritta la quantizzazione e codifica vettoriale. Nella seconda parte dell'articolo vengono esposti i risultati ottenuti applicando tali tecniche all'algoritmo di DCT Ibrida descritto nella bibliografia citata precedentemente.

Il lavoro illustrato è stato condotto presso il *Centro Ricerche* RAI nell'ambito del progetto Finalizzato Telecomunicazioni del CNR, Sottoprogetto 3, Tema 3.4 «Studi sull'HDTV».

2. Generalità sulla codifica

I simboli emessi da una sorgente possono essere codificati singolarmente oppure raggruppati in blocchi prima della loro codifica. Si ottiene in tal modo una codifica vettoriale in quanto si formano vettori di dimensione k e ad ognuno di essi si associa un simbolo. Supponendo che ogni elemento del vettore possa assumere uno degli M valori

* Ing. Laurent Boch del Politecnico di Torino, dott. Mario Stroppiana del Centro Ricerche RAI - Torino.
Dattiloscritto pervenuto alla redazione il 18 aprile 1991.
Il lavoro è stato svolto nell'ambito del Progetto Finalizzato Telecomunicazioni del CNR.
1 Al progetto EU-256 partecipano il Centro Ricerche RAI, la Telettra Italiana, la Retevision, la Telettra Spagnola e l'Università Politecnica di Madrid.

che costituiscono l'alfabeto della sorgente, i vettori di dimensione k potranno assumere uno degli M^k valori costituenti l'alfabeto della sorgente che emette vettori. La codifica a lunghezza variabile dell'alfabeto della sorgente scalare o di quella vettoriale consiste nell'assegnare, ad ogni simbolo o vettore generato, una parola di codice la cui lunghezza, espressa in numero di bit, è inversamente proporzionale alla sua probabilità (bibl. 4).

2.1 ENTROPIA DI SORGENTE

Si consideri una sorgente discreta X che genera una sequenza x_0, x_1, ..., x_n, ... di simboli che possono assumere ciascuno uno qualsiasi degli M valori a_j che costituiscono l'alfabeto A della sorgente.

Si indica con $I(a_j)$, l'*Informazione* di un generico valore a_j, appartenente all'alfabeto della sorgente discreta che risulta, espressa in bit (bibl. 7):

[1] $$I(a_j) = \log_2(1/p(a_j))$$

Il contenuto d'informazione media $H(x_n)$ del generico simbolo x_n è uguale all'entropia H(A) dell'alfabeto della sorgente e risulta:

[2] $$H(x_n) = H(A) = \sum_{1}^{M}{}_j \; p(a_j) \cdot I(a_j)$$

Se l'informazione è data dalla [1], ossia è espressa in bit, allora l'entropia rappresenta il minimo numero medio di bit necessari a codificare i simboli dell'alfabeto A e tale valore soddisfa la relazione:

[3] $$0 \leq H(A) \leq \log_2(M)$$

dove $\log_2$ (M) indica l'entropia di un alfabeto in cui i simboli a_j hanno tutti la stessa probabilità $p(a_j) = 1/M$ di occorrenza.

Il contenuto medio d'informazione del generico simbolo x_n sarà in generale inferiore se si suppone la conoscenza del simbolo emesso dalla sorgente all'istante precedente $n - 1$. L'entropia condizionata $H(x_n | x_{n-1})$ risulta:

[4] $$H(x_n | x_{n-1}) = -\sum \sum P(x_n, x_{n-1}) \cdot \log_2(P(x_n | x_{n-1}))$$

$H(x_n | x_{n-1})$ risulta minore od uguale all'entropia non condizionata $H(x_n)$; l'eguaglianza si verifica solo nel caso in cui i simboli emessi dalla sorgente sono indipendenti tra loro. L'entropia condizionata del simbolo x_n risulterà inoltre ancora minore se si suppone la conoscenza non solo del simbolo precedente x_{n-1}, ma anche del simbolo x_{n-2} e così via. Il suo valore è ottenuto, in tale caso, mediante una semplice estensione della [4].

Da quanto esposto sembra conveniente, in generale, raggruppare più simboli emessi dalla sorgente, ossia considerare per esempio l'entropia della coppia (x_n, x_{n-1}) che risulta:

[5] $$H(x_n, x_{n-1}) = H(x_{n-1}) + H(x_n | x_{n-1}) \leq \leq H(x_n) + H(x_{n-1})$$

ed è quindi, in generale, inferiore alla somma dell'entropia dei singoli simboli[2]. È facilmente intuibile che il guadagno aumenta se si raggruppano più simboli; nel caso in cui tutti i simboli emessi dalla sorgente siano raggruppati in un unico vettore si ottiene un'entropia pari a quella di sorgente. La [5] rappresenta il guadagno in termini di riduzione di entropia della codifica vettoriale rispetto a quella scalare; tale guadagno viene chiamato *guadagno di memoria*.

2.2 CODIFICA DI SORGENTE

Si supponga di voler trasmettere i simboli di una sorgente su un canale binario.

Ad ogni simbolo deve essere associato un numero intero di bit. La *lunghezza media* $\tilde{n}$ delle parole di codice risulta quindi:

[6] $$\tilde{n} = \sum_{1\,j}^{M} p(a_j) \cdot n_j$$

dove M indica il numero di simboli dell'alfabeto ed n_j il numero di bit della parola di codice associata al simbolo a_j dell'alfabeto.

Si può dimostrare che esiste almeno un codice (bibl. 7), soddisfacente la *regola del prefisso*, tale che:

[7] $$H(x) \leq \tilde{n} \leq H(x) + 1$$

La lunghezza media $\tilde{n}$ risulta uguale all'entropia dell'alfabeto solo nel caso in cui la probabilità di ogni simbolo sia data da:

[8] $$p(a_j) = 2^{-n_j}$$

Se si raggruppano k simboli emessi dalla sorgente e si codifica il vettore di dimensione k risultante, allora ad ogni vettore deve essere associato un simbolo costituito da un numero intero di bit; il numero medio di bit per vettore risulta:

[9] $$\tilde{n}_v = \sum_{1\,j1}^{M} \cdots \sum_{1\,jk}^{M} p(a_{j1}, \ldots, a_{jk}) \cdot n_{j1, \ldots jk}$$

dove M indica il numero di simboli dell'alfabeto A ed $n_{j1}, \ldots, {}_{jk}$ il numero di bit della parola di codice associata al vettore $(a_{j1}, \ldots, a_{jk})$. Il numero medio di bit per vettore $\tilde{n}_v$, risulta compreso tra i limiti definiti dalla [7]; supponendo indipendenti tra loro i simboli emessi dalla sorgente, il numero medio di bit per simbolo soddisfa la relazione:

[10] $$H(x) \leq \tilde{n} = \frac{\tilde{n}_v}{k} \leq H(x) + \frac{1}{k}$$

ossia rimane compreso tra l'entropia dell'alfabeto della sorgente e l'entropia + $1/k$ dove k indica la dimensione del vettore. Il risultato espresso dalla [10] è indicato come *guadagno di codifica* vettoriale.

Supponiamo di codificare vettorialmente un vettore v di dimensione k i cui elementi possono assumere M valori, il numero dei possibili valori assunti dal vettore è M^k. La dimensione dell'alfabeto per la codifica vettoriale cresce esponenzialmente con la dimensione dell'alfabeto;

---

2 Si ricorda che $H(x_n, x_{n-1}) = H(x_{n-1}) + H(x_n)$ solo nel caso in cui i due simboli sono indipendenti tra loro.

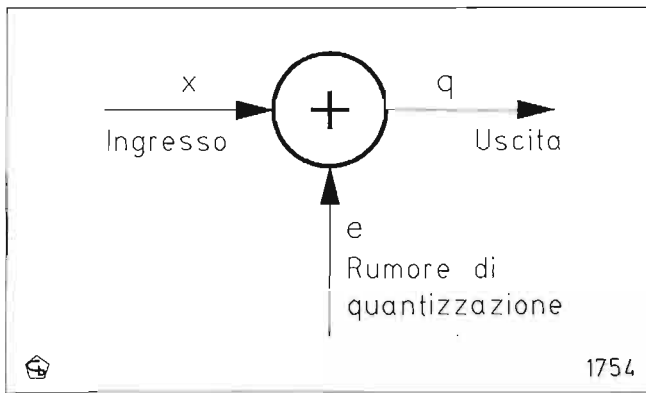



Fig. 1 — Modello additivo di un quantizzatore. L'errore di quantizzazione **e** è dato da: **e** = **q** - **x**, dove **q** indica il segnale quantizzato e **x** il segnale all'ingresso del quantizzatore.

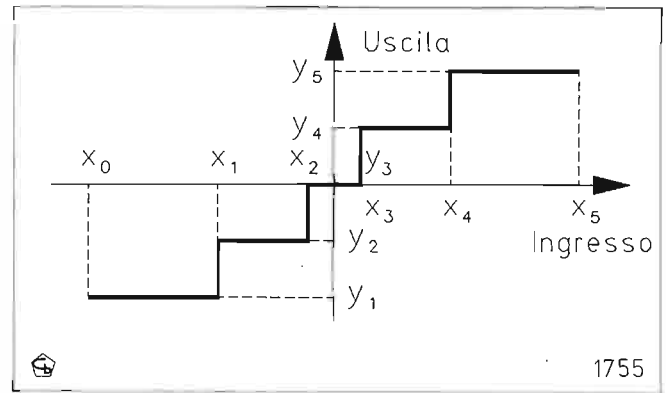



Fig. 2 — Caratteristica di un quantizzatore non uniforme con 5 livelli d'uscita, $N = 5$. Con $x_0, x_1, \ldots, x_5$ si sono indicate le soglie di separazione tra gli intervalli di quantizzazione e con $y_1, y_2 \ldots, y_5$ i valori d'uscita o livelli di restituzione del quantizzatore. Il quantizzatore ottimo fissato il numero di livelli N di uscita risulta quello che minimizza l'errore di quantizzazione e, dato da:

$$e = \sum_{i=1}^{N} \int_{x_{i-1}}^{x_i} (x - y_i)^2 \cdot p(x)dx$$

dove $p(x)$ indica la densità di distribuzione del segnale d'ingresso.

per applicazioni sul segnale video non codificato[3] si ha M = 256, quindi M^k risulta molto grande già nel caso $k = 2$.

3. Generalità sulla quantizzazione

Il processo di quantizzazione sostituisce campioni analogici con valori approssimati presi da un insieme finito di valori permessi. Il segnale campionato **q** consiste quindi di due componenti: il valore del segnale analogico d'ingresso **x**, che è il segnale utile, più il valore dell'errore di quantizzazione **e** che è il segnale di rumore, figura 1. La precisione del processo di quantizzazione risulta definita dal valore quadratico medio dell'errore di quantizzazione che deve quindi essere minimizzato.

Il quantizzatore può essere di tipo scalare oppure di tipo vettoriale.

3.1 QUANTIZZAZIONE SCALARE

Il quantizzatore scalare a N punti è definito specificando un insieme di $N + 1$ livelli $x_0, x_1, \ldots, x_N$ che definiscono gli intervalli in cui è spezzata la caratteristica di quantizzazione ed N valori d'uscita $y_1, y_2, \ldots, y_N$ associati a ciascun intervallo, come illustrato nelle figure 2 e 3. Gli intervalli del quantizzatore possono essere diversi tra loro, quantizzatore non uniforme, figura 2, oppure tutti uguali, quantizzatore uniforme, figura 3.

Il quantizzatore che, fissato il numero di livelli N di quantizzazione, minimizza il valore quadratico medio dell'errore di quantizzazione è detto quantizzatore «ottimo nel senso di Max» (bibl. 8) o secondo la «constrained resolution» ed è illustrato in figura 2.

L'ampiezza di ogni suo intervallo di quantizzazione è inversamente proporzionale alla probabilità di occorrenza del segnale da quantizzare all'interno dello stesso. la densità di distribuzione del segnale quantizzato e pressoché uniforme e quindi tale segnale può essere codificato con codici a lunghezza fissa (FLC, Fixed Lenght Code), dove la lunghezza in bit delle parole è determinata dal numero di livelli d'uscita del quantizzatore.

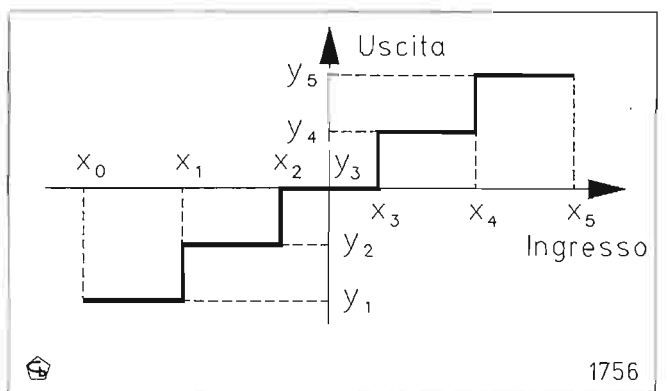



Fig. 3 — Caratteristica di un quantizzatore uniforme. La densità di distribuzione del segnale di uscita risulta non uniforme, per cui le prestazioni di tale quantizzatore devono essere definite dall'errore di quantizzazione e dato da:

$$e = \sum_{i=1}^{N} \int_{x_0+(i-1)\cdot\delta x}^{x_0+i\cdot\delta x} (x - y_i)^2 \cdot p(x)dx$$

dove δx indica l'intervallo costante tra le soglie del quantizzatore, e dall'entropia **r** del segnale quantizzato data da:

$$r = -\Sigma p_i \cdot \log_2 p_i$$

Con p_i si è indicata la probabilità del livello i-esimo del segnale quantizzato; tale probabilità risulta data da

$$p_i = \int_{x_{i-1}}^{x_i} p(x)dx$$

Le prestazione del quantizzatore uniforme sono molto vicine a quelle del quantizzatore ottimo con la condizione di «constrained entropy», ossia minimizzazione dell'errore di quantizzazione a parità di entropia del segnale quantizzato.

3 Si fa presente che la precisione dei coefficienti DCT del segnale video codificato mediante DPCM o trasformate è superiore ad 8 bit, 12 bit nello schema DCT definito dal CMTT/2, quindi in tale caso il valore di M risulta superiore.

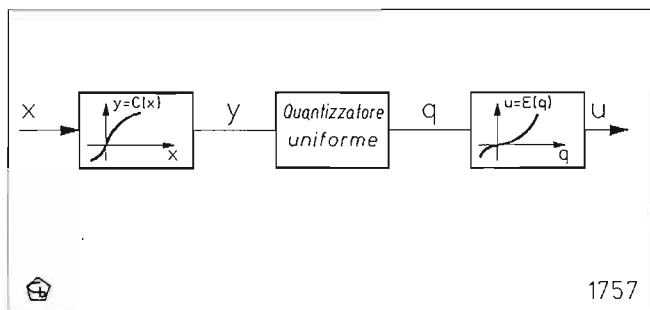



Fig. 4 — Quantizzatore non uniforme ottenuto ponendo un sistema non lineare prima e dopo il quantizzatore uniforme. Le condizioni di minimo errore di quantizzazione, quantizzatore ottimo, si hanno con:

$$C(x) = c \cdot \int_0^x [p(s)]^{1/3}\, ds, \quad E(x) = C^{-1}(x)$$

dove $p(x)$ indica la densità di distribuzione del segnale d'ingresso al compressore, $C(x)$ indica la caratteristica ingresso/uscita di compressione ed $E(x)$ indica la caratteristica ingresso/uscita dell'espansore.

Se il segnale da quantizzare ha una densità di distribuzione uniforme il quantizzatore ottimo nel senso di Max risulta quello uniforme. Tale constatazione permette la progettazione del quantizzatore ottimo mediante un compressore che rende uniforme la densità di distribuzione del segnale, un quantizzatore uniforme ed un espansore che ripristina la densità di distribuzione originaria del segnale, figura 4 (bibl. 9).

Il segnale in uscita dal quantizzatore uniforme ha una distribuzione statistica non uniforme il cui andamento è uguale a quello del segnale d'ingresso. Una codifica a lunghezza fissa del segnale quantizzato risulta, in questo caso, altamente inefficiente. Bisogna quindi utilizzare codici a lunghezza variabile (VLC, Variable Length Code) che assegnano ai vari livelli d'uscita del quantizzatore parole di codice la cui lunghezza è inversamente proporzionale alla probabilità di occorrenza del livello stesso (bibl. 4). In tale caso il quantizzatore ottimo risulta quello che minimizza l'errore di quantizzazione a parità di entropia del segnale quantizzato, come illustrato in Fig. 3; tale quantizzatore viene detto ottimo sotto le condizioni della «constrained entropy» (bibl. 10). Le operazioni di quantizzazione e codifica, seppur distinte tra loro, devono essere considerate congiuntamente nel processo di ottimizzazione.

Un sistema di codifica per la trasmissione su canali a bit-rate costante varia la precisione del quantizzatore in funzione del riempimento del buffer di canale (bibl. 2 e 11). Tale buffer viene scritto a velocità variabile e letto ad una velocità costante determinata dalla capacità del canale. In tale situazione il quantizzatore uniforme risulta molto prossimo al quantizzatore ottimo e comunque migliore di quello non uniforme, ottimo nel senso di Max (bibl. 9, 11 e 12).

Nel caso in cui si conosca la densità di distribuzione del segnale da quantizzare si possono ulteriormente migliorare le prestazioni del quantizzatore uniforme ponendo il valore di restituzione di ogni intervallo uguale al corrispettivo baricentro. Tale valore minimizza la distorsione del quantizzatore, senza modificare l'entropia del segnale quantizzato (bibl. 13).

I simboli in uscita dal quantizzatore scalare possono essere codificati scalarmente, figura 5a, oppure raggruppati in vettori e codificati vettorialmente, figura 5b; in questo secondo caso si ottengono i guadagni di memoria e di codifica rispetto alla codifica scalare.

Fig. 5 — Quantizzazione scalare con codifica scalare *(a)* e con codifica vettoriale *(b)*.
a — ciascun simbolo x_i, è quantizzato scalarmente dal quantizzatore QS ottenendo il simbolo q_i che viene poi, mediante CS, codificato scalarmente a lunghezza variabile, VLC, od a lunghezza fissa, FLC, ottenendo il simbolo c_i;
b — i simboli emessi dalla sorgente, x_i, vengono quantizzati scalarmente dal quantizzatore QS ottenendo i simboli q_i che sono poi raggruppati, dal modulo VETT, nel generico vettore $\mathbf{q}_i(q_1, \ldots, q_k)$ di k elementi; ad ogni vettore $\mathbf{q}_i$ è associato, mediante il codificatore vettoriale CV, un simbolo v_i che viene poi, tramite il codificatore scalare CS, codificato nel simbolo c_i.

3.2 QUANTIZZAZIONE VETTORIALE

La quantizzazione scalare è un'operazione che assegna, al singolo campione di un segnale analogico, un valore, scelto da un insieme finito, che meglio approssima il valore reale del campione del segnale.

La quantizzazione vettoriale, invece, assegna un valore, scelto da un insieme finito di valori rappresentativi, ad una successione di campioni del segnale analogico. Tale operazione trasforma un vettore in un simbolo scalare e quindi comprende anche parte delle operazioni descritte nel capitolo 2 come codifica vettoriale. Il valore assegnato al vettore può essere codificato con parole a lunghezza fissa od a lunghezza variabile in dipendenza della densità

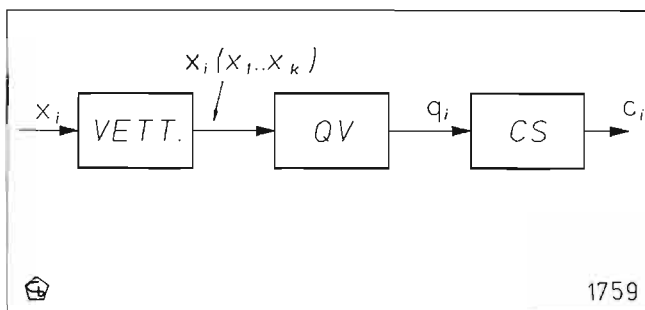



Fig. 6 — Quantizzazione vettoriale. I simboli emessi dalla sorgente, x_i, vengono raggruppati da VETT in vettori $x_i(x_1, \ldots, x_k)$ di k elementi, ogni vettore x_i è quantizzato vettorialmente dal modulo QV che associa ad esso un simbolo q_i che viene poi codificato dal codificatore scalare CS nel simbolo c_i.

di distribuzione dei simboli in uscita dal quantizzatore. In figura 6 è riportato lo schema a blocchi relativo alla quantizzazione e codifica vettoriale.

Il ricevitore associa al simbolo ricevuto e decodificato un vettore che meglio approssima il vettore originario presente all'ingresso del quantizzatore vettoriale. La precisione della ricostruzione dipende dalla dimensione dell'insieme dei valori rappresentativi, detto vocabolario «code book».

Anche la quantizzazione vettoriale può essere di tipo uniforme oppure non uniforme.

3.2.1 *Vantaggi della quantizzazione vettoriale*

I guadagni della quantizzazione vettoriale rispetto a quella scalare sono di due tipi: *guadagno di riempimento dello spazio di quantizzazione* e *guadagno di forma della densità di distribuzione del segnale da codificare* (bibl. 14 e 15).

Tali guadagni sono quelli che presenta lo schema di quantizzazione e codifica illustrato in figura 6 rispetto a quello di figura 5b; nei confronti dello schema di quantizzazione e codifica scalare di figura 5a si hanno ancora i guadagni propri della codifica vettoriale e quindi quello di codifica e quello di memoria, quest'ultimo presente solo quando la sorgente emette simboli correlati.

— *Guadagno di riempimento dello spazio di quantizzazione.* L'operazione di quantizzazione vettoriale con vettori di dimensione k è fatta in uno spazio a k dimensioni; ogni componente del vettore è associata ad una dimensione dello spazio. Se ogni elemento del vettore è quantizzato indipendentemente dagli altri, ovvero si ha una quantizzazione scalare, lo spazio di quantizzazione viene diviso in celle ipercubiche, detti k-cubi. Utilizzando la quantizzazione vettoriale, ovvero l'associazione di un simbolo di riproduzione per ogni vettore, si ha invece un'ampia libertà nella scelta della cella multidimensionale e quindi si può riempire lo spazio

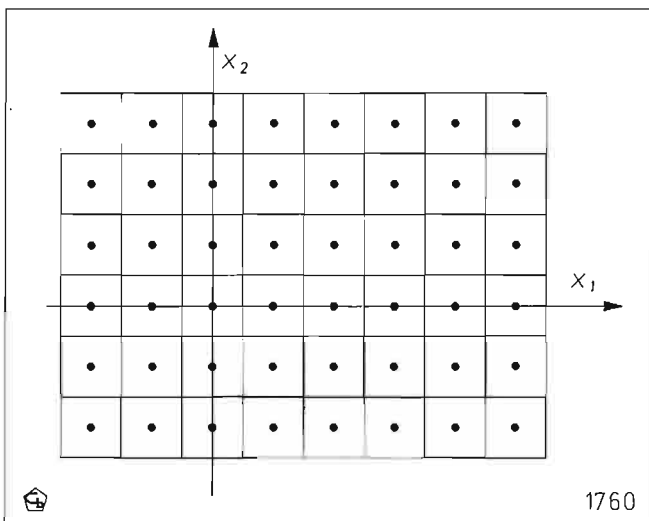



Fig. 7 — Quantizzatore vettoriale bidimensionale uniforme con celle quadrate. Tale divisione del piano di quantizzazione può essere ottenuta mediante quantizzazione scalare delle singole componenti del vettore.

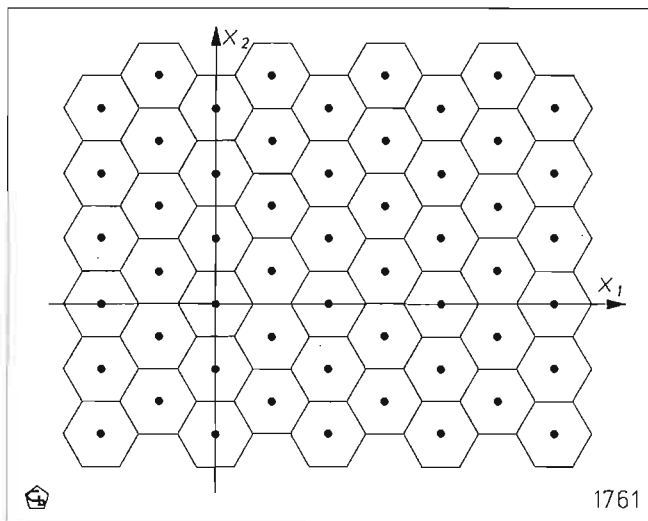



Fig. 8 - Quantizzazione vettoriale uniforme con vettori a due elementi. La ricopertura del piano di quantizzazione avviene mediante celle esagonali che risultano ottime in quanto minimizzano la distorsione del segnale quantizzato a parità di area delle celle di quantizzazione. La distorsione del segnale quantizzato risulta uguale a 0,962 volte quella fornita da un quantizzatore a celle quadrate.

di quantizzazione mediante celle che forniscono una distorsione inferiore rispetto alla cella ipercubica. Quanto esposto è illustrato nelle figure 7 e 8 nel caso di uno spazio di quantizzazione bidimensionale, $k = 2$, e con distribuzione uniforme del segnale da quantizzare; la divisione dello spazio di quantizzazione con celle esagonali fornisce, a parità di area della cella e quindi a parità di entropia quando il segnale d'ingresso ha una distribuzione uniforme, un errore di quantizzazione uguale a 0,962 volte quello ottenuto con le celle quadrate.

— *Guadagno di forma della densità di distribuzione.* Al fine di minimizzare l'errore di quantizzazione è conveniente addensare le celle di quantizzazione nelle zone in cui la densità di distribuzione del segnale da quantizzare è elevata ed invece avere celle meno addensate nelle zone in cui la densità di distribuzione è bassa. Un quantizzatore scalare opera indipendentemente su ogni elemento del vettore e quindi può solo addensare le proiezioni delle celle sui singoli assi dello spazio di quantizzazione in base alla densità di distribuzione marginale del segnale da quantizzare. La cella di quantizzazione risulta data dal prodotto delle sue proiezioni lungo gli assi e quindi, con distribuzioni del segnale diverse da quella uniforme, il suo addensamento non sarà più conforme alla densità di distribuzione del segnale.

3.2.2 *Quantizzazione vettoriale uniforme e non uniforme*

La quantizzazione vettoriale può essere uniforme, quando tutte le celle di quantizzazione sono uguali tra loro oppure non uniforme quando le celle sono tra loro diverse.

Il quantizzatore vettoriale ottimo non uniforme, figura 9, risulta un'estensione del corrispettivo scalare; esso definisce la dimensione delle celle in modo tale da avere,

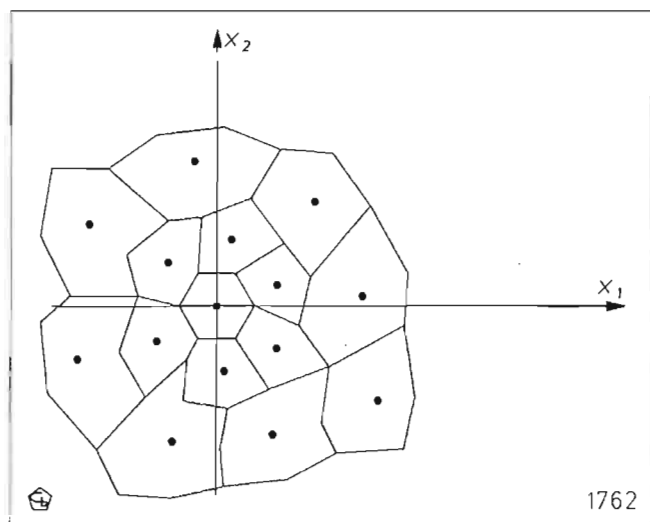



Fig. 9 — Quantizzatore vettoriale bidimensionale non uniforme. La dimensione delle celle di quantizzazione dipende dalla densità di distribuzione del segnale da quantizzare. Le celle del quantizzatore ottimo rendono infatti uniforme la densità di distribuzione del segnale quantizzato rappresentato dal centroide di ogni cella.

fissato il numero delle stesse, la minima distorsione sul segnale quantizzato, «constrained resolution». Tale quantizzatore ha, rispetto al corrispettivo quantizzatore scalare, tutti i guadagni esposti precedentemente. Esso non presenta il guadagno di forma solo quando la distribuzione del segnale d'ingresso è uniforme (bibl. 15); in tale caso il quantizzatore ottimo risulta quello uniforme. Uno svantaggio presentato da tale tipo di quantizzatore è legato alla quantità di memoria richiesta dall'insieme dei simboli di quantizzazione, «code book», ed alla complessità del processo di codifica. I simboli di quantizzazione, detti centroidi della cella, coincidono, al fine di minimizzare la distorsione, con il baricentro della cella stessa (bibl. 15). La densità di distribuzione dei simboli all'uscita del quantizzatore risulta pressoché uniforme e quindi tali simboli sono codificati con parole a lunghezza fissa definita dal numero di centroidi del quantizzatore.

Il quantizzatore uniforme, figura 8, ha tutte le celle di quantizzazione uguali tra loro ed i centroidi delle celle costituiscono un reticolo (bibl. 16). Tale quantizzatore non richiede quindi la memoriazzazione del code book ed inoltre l'operazione di quantizzazione risulta semplice.

A parità di entropia del segnale quantizzato, il quantizzatore che minimizza la distorsione, quantizzatore ottimo nella condizione di «constrained entropy», può essere determinato mediante complessi processi di ottimizzazione e risulta comunque molto prossimo al quantizzatore uniforme (bibl. 17). Tale situazione è quella che si presenta, come già detto, nei casi di codifica entropica, ossia quando si ha un buffer di canale che retroagisce sul quantizzatore. In tale condizione il quantizzatore vettoriale presenta, rispetto a quello scalare, i guadagni sopra esposti eccettuato il guadagno di forma della densità di distribuzione che risulta uguale a 1 (bibl. 15). La densità di distribuzione del segnale quantizzato risulta non uniforme.

4. Tipi di quantizzatori vettoriali

4.1 LINDE-BUZO-GRAY

Il metodo noto come Linde-Buzo-Gray (LBG) ha lo scopo di progettare un quantizzatore vettoriale ottimo nel senso che minimizza l'errore di quantizzazione mantenendo fisso il numero di celle di quantizzazione e quindi il numero di simboli in uscita dal quantizzatore stesso (bibl. 18 e 19). La densità di distribuzione dei campioni del segnale quantizzato è uniforme e quindi tali campioni vengono codificati con parole a lunghezza fissa.

Il quantizzatore ottimo deve essere progettato sulla base della densità di distribuzione dei vettori da quantizzare; in genere, tale distribuzione non è perfettamente nota od i modelli matematici che la rappresentano non sono sufficientemente aderenti, od ancora essa è troppo complicata per sfruttarne l'analisi in sede di progetto. La progettazione del quantizzatore viene quindi fatta utilizzando una sequenza di prova con caratteristiche statistiche simili a quelle delle sequenze su cui opererà il quantizzatore ed utilizzando un vocabolario, code book, iniziale che contiene i simboli d'uscita del quantizzatore. L'estensione del vocabolario dipende dal bit rate richiesto. La fase di progettazione consiste nella determinazione ottima, sulla base del minimo valore quadratico medio dell'errore di quantizzazione, delle celle di ricopertura dello spazio di quantizzazione e nella determinazione ottima del simbolo associato ad ogni cella, detto centroide o baricentro della cella. Le due fasi di ottimizzazione sono eseguite in modo iterativo ed il processo di progettazione si ferma quando l'errore di quantizzazione risulta minore di quello prefissato (bibl. 18).

È evidente che l'algoritmo tende ad adattare la forma e la dimensione delle celle alla distribuzione dei vettori d'ingresso; dove questi sono più addensati anche le riproduzioni sono addensate e le rispettive celle sono più piccole. In questo modo la distribuzione dei simboli in uscita dal quantizzatore tende ad essere uniforme e risulta minimizzata la distorsione media.

Il quantizzatore progettato con questo metodo presenta alcuni inconvenienti. I simboli di riproduzione non risultano in generale ordinati secondo uno schema particolare e quindi è necessario utilizzare un vocabolario che li contenga; l'assegnazione del simbolo di riproduzione al vettore d'ingresso deve essere fatta sulla base della minima distorsione e quindi tale distorsione deve essere calcolata tra il vettore d'ingresso e tutti i simboli di riproduzione contenuti nel vocabolario (metodo di «*ricerca completa*»); ad ogni bit rate deve essere associato un proprio vocabolario, ottimizzato, dei simboli di riproduzione.

I precedenti inconvenienti possono essere, in parte, ridotti od eliminati progettando ed utilizzando quantizzatori vettoriali sub-ottimi. Con tali quantizzatori il simbolo riproduttore del vettore da quantizzare non sarà più determinato mediante il metodo di ricerca completa su tutto il vocabolario, ma mediante altri metodi, tipo ricerca ad albero oppure quantizzazione multistadio con vocabolari ridotti.

4.1.1 *Soluzioni alternative alla ricerca completa*

Vi sono una varietà di quantizzatori vettoriali pensati

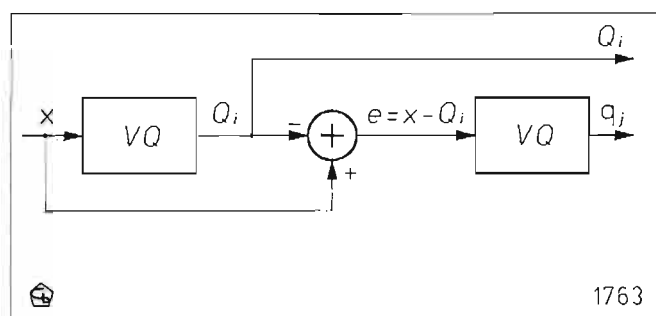



Fig. 10 — Quantizzatore vettoriale a 2 stadi. Il primo stadio vettoriale quantizza grossolanamente il segnale d'ingresso, mentre il secondo stadio quantizza finemente l'errore di quantizzazione del primo stadio. Si richiede la memorizzazione di due differenti vocabolari, uno per ogni stadio di quantizzazione vettoriale.

per rimuovere gli inconvenienti della «ricerca completa» sul quantizzatore tipo Linde-Buso-Gray, con una leggera perdita di efficienza (bibl. 19).

Tra i quantizzatori vettoriali senza memoria sono particolarmente interessanti i casi di:

— *quantizzatore vettoriale con ricerca ad albero*. Il vocabolario è progettato in modo da dimezzare ad ogni calcolo di distorsione il numero di riproduzioni possibili; sostanzialmente l'intero alfabeto è stato generato da una successione di sdoppiamenti a partire da un unico punto; se la profondità dell'albero è R, ovvero si devono prendere R decisioni per arrivare al punto finale, l'intero vocabolario è formato da 2^R punti con un evidente risparmio in calcolo rispetto alla ricerca completa. Purtroppo raddoppia il vocabolario che deve essere memorizzato perché bisogna conservare i simboli di riproduzione ad ogni livello di profondità. È possibile regolare il bit rate a passi di un bit per vettore e quindi a passi di $1/k$ bit per simbolo, dove k indica la dimensione del vettore, utilizzando vocabolari ai vari livelli dell'albero.

— *quantizzatore vettoriale multistadio*. Un quantizzatore vettoriale a 2 stadi è illustrato in figura 10. Il primo stadio codifica in modo grossolano il vettore d'ingresso mentre il secondo stadio codifica, con passo più fine, l'errore commesso dal primo quantizzatore. In tale modo si riesce a ridurre sia la dimensione del vocabolario sia il numero di calcoli della distorsione. La quantizzazione multistadio può essere associata anche a quantizzatori vettoriali diversi dal Linde-Buzo-Gray.

4.2 Quantizzatore piramidale

Il quantizzatore vettoriale piramidale ha il pregio di essere basato su considerazioni teoriche che non lo limitano eccessivamente in quanto a praticità per cui si può adattare a diversi valori del bit rate (bibl. 20 e 21).

Si osserva che la densità di distribuzione di una sorgente può essere, in generale, non uniforme ed inoltre può presentare una zona ristretta in cui essa è molto più grande che altrove; un tale comportamento si riscontra ad esempio sui coefficienti DCT di un'immagine. Nel caso di sorgenti vettoriali in cui tutti gli elementi costituenti il vettore sono indipendenti tra loro ed hanno uguale distribuzione statistica, i vettori da quantizzare si addensano attorno alla zona di massimo della densità di distribuzione e tale addensamento aumenta all'aumentare della dimensione del vettore. Tale regione assume inoltre forme regolari definite dal tipo di distribuzione degli elementi costituenti il vettore, ossia dei simboli emessi dalla sorgente; nel caso in cui tale distribuzione sia una *laplaciana*[4] essa risulta la superficie di una *iperpiramide* avente un numero di dimensioni pari al numero di elementi che costituiscono il vettore.

Ciò significa che, supponendo di essere nel caso di simboli indipendenti e identicamente distribuiti e con dimensione del vettore grande, si introduce una distorsione trascurabile se si sostituisce al vettore d'ingresso il più prossimo vettore giacente sulla superficie della piramide in questione.

Si può allora scegliere come vocabolario dei simboli di riproduzione il sottoinsieme di punti di un reticolo (ad es. ipercubico) che interseca la piramide. I punti di tale reticolo risultano i simboli d'uscita del quantizzatore ed hanno una densità di distribuzione quasi uniforme; essi possono essere codificati con parole a lunghezza fissa. Inoltre le operazioni di co-decodifica possono trarre vantaggio dalla struttura a reticolo del quantizzatore e quindi risultare più semplici delle rispettive operazioni richieste dal quantizzatore LBG (Linde-Buzo-Gray).

I simboli emessi dalla sorgente possono essere correlati e non avere la stessa varianza. Nel primo caso è necessario scorrelare tra loro tali simboli mediante l'applicazione di tecniche predittive od a trasformate, mentre nel secondo caso bisogna uniformare le varianze dei simboli mediante coefficienti moltiplicativi, detti *pesi*.

I *pesi* e la distorsione accettata sul segnale quantizzato definiscono la dimensione della piramide e quindi il numero dei simboli d'uscita del quantizzatore vettoriale e di conseguenza, essendo la codifica a lunghezza fissa, il bit rate.

Così, mentre dalla natura strutturata della distribuzione dei vettori d'uscita dipende la relativamente bassa complessità di questo quantizzatore, dalla natura parametrica deriva la possibilità di un suo utilizzo adattativo e di una discreta regolazione del bit rate.

Questo tipo di quantizzatore vettoriale è stato usato nella codifica del segnale televisivo mediante DCT (bibl. 23).

4.3 Quantizzatore a reticolo

In pratica si tratta del quantizzatore vettoriale uniforme. Come si è già visto, data la dimensione del vettore si può scegliere tra diverse forme di celle in grado di riempire lo spazio di quantizzazione e una di queste celle è proprio l'ipercubo, cella utilizzata nel caso della quantizzazione scalare.

La scelta della cella base del quantizzatore è fatta considerando da una parte il suo guadagno rispetto alla cella ipercubica relativa al caso scalare e dall'altra l'esistenza di algoritmi che permettano una semplice e veloce opera-

4 I coefficienti DCT di un'immagine hanno appunto una densità di distribuzione laplaciana (bibl. 22).

zione di quantizzazione (bibl. 24, 25).

Il guadagno ottenibile al variare della forma della cella è detto guadagno di riempimento dello spazio ed è massimo con le celle ipersferiche che però non danno una copertura completa dello spazio. Supponendo che i simboli siano scorrelati tra loro ed ugualmente distribuiti, è possibile definire, per ogni tipo di cella, il suo guadagno di riempimento dello spazio rispetto alla cella ipercubica (bibl. 15); le celle di quantizzazione devono inoltre essere molto piccole, in modo tale che si possa considerare uniforme la distribuzione del segnale all'interno di ogni cella di quantizzazione. Le ipotesi precedenti (quantizzatore con celle di quantizzazione molto piccole e quindi con un elevato numero di centroidi) sono alla base della teoria dell'«high resolution» e sono valide quando si opera ad elevati bit rate.

Nel caso di quantizzazione vettoriale con vettori di lunghezza 2, 3 e 4 è possibile determinare la cella ottima di quantizzazione che risulta esagonale nel caso in cui la lunghezza del vettore è uguale a 2 ed a forma di octaedro troncato con vettori di dimensione 3 e di dimensione 4.

Quando il quantizzatore lavora a bassa risoluzione non si può più considerare uniforme la distribuzione del segnale all'interno delle celle di quantizzazione e quindi non si possono più utilizzare i risultati della teoria dell'«high resolution». In tale caso, come detto nel paragrafo 3.2.2, si può ottimizzare il quantizzatore mediante processi iterativi che considerano anche l'operazione di codifica; il quantizzatore ottenuto con tale ottimizzazione risulta comunque molto prossimo a quello uniforme, ossia quello con tutte le celle di quantizzazione uguali tra loro.

5. Applicazione della codifica vettoriale all'algoritmo DCT

La codifica DCT del segnale televisivo può essere fatta secondo lo schema detto di *bit allocation* oppure quello detto di *codifica entropica.*

Si fornisce una breve descrizione dei due metodi con particolare attenzione all'applicazione della quantizzazione vettoriale.

5.1 DCT CON «BIT ALLOCATION»

Secondo la teoria della *Rate Distorsion* (bibl. 26) i simboli emessi da una sorgente possono essere codificati con un numero medio di bit R, che dipende dalla varianza σ^2 della sorgente e dalla distorsione massima tollerata D, secondo la formula:

$$R = \frac{1}{2} \cdot \log_2 \frac{\sigma^2}{D} \qquad [11]$$

I coefficienti DCT hanno una varianza che dipende dalla loro posizione all'interno del blocco di codifica (bibl. 27 e 22) ed inoltre anche la distorsione tollerata può variare al variare della posizione del coefficiente all'interno del blocco. La formula [11] definisce allora un numero di bit diverso da coefficiente a coefficiente. Si può quantizzare ogni coefficiente DCT con un quantizzatore ottimo avente un numero di livelli definito da R e codificare poi l'uscita di tale quantizzatore con simboli la cui lunghezza è fissa ed uguale a R bit (bibl. 28).

Purtroppo i valori forniti dalla [11] non saranno in generale interi per cui diventa necessario quantizzare e codificare il singolo coefficiente scegliendo tra il valore intero superiore, che provoca un incremento del bit rate, e quello intero inferiore, che aumenta il valore prefissato della distorsione. Così facendo ci si allontana molto dalle condizioni ottime di codifica stabilite dalla *Rate Distorsion.*

Per avvicinarsi alle condizioni ottimali di codifica è necessario poter codificare il singolo coefficiente con un numero medio di bit frazionario. Questo è possibile formando vettori i cui elementi hanno tutti la stessa varianza e quindi lo stesso valore di R definito dalla [11] ed eseguendo quindi una quantizzazione vettoriale tale che i simboli d'uscita dal quantizzatore abbiamo una densità di distribuzione uniforme. In tal modo si codifica il vettore con un numero intero di bit, che può però risultare un valore frazionario per i singoli coefficienti.

Può costituire un problema la formazione di vettori i cui elementi abbiano la stessa varianza; le soluzioni adottabili sono la codifica vettoriale di coefficienti appartenenti a blocchi diversi, ma con la stessa posizione nel blocco, oppure la formazione di vettori secondo le diagonali secondarie del blocco DCT (bibl. 29) che sono grosso modo il luogo dei coefficienti a varianza costante.

In tale schema di codifica, oltre al vantaggio appena descritto, la quantizzazione vettoriale fornisce anche i vantaggi della codifica e della quantizzazione vettoriale descritti nei paragrafi 2.2 e 3.2.

D'altra parte l'utilizzo del quantizzatore ottimo non uniforme comporta gli inconvenienti realizzativi descritti precedentemente, specialmente nel caso vettoriale.

5.2 ALGORITMO DCT CON CODIFICA ENTROPICA E BUFFER DI TRASMISSIONE

Uno schema di codifica dei coefficienti DCT utilizzante un quantizzatore scalare uniforme, la codifica entropica e la regolazione dei bit rate per mezzo di un buffer di trasmissione è già stato oggetto di precedenti pubblicazioni (bibl. 2 e 4).

Sostanzialmente il flusso di bit, la cui costanza non può essere garantita dalla codifica a lunghezza variabile, è regolato dal buffer di trasmissione che, non solo fa da tampone per l'uscita sul canale, ma agisce anche in retroazione sul quantizzatore modificando il fattore di scalamento dei coefficienti. In questo modo la regolazione del bit rate può essere fine a piacere consentendo allo schema di fornire delle buone prestazioni con una complessità del tutto accettabile. La codifica a lunghezza variabile dei coefficienti DCT unita alla codifica delle sequenze di zeri e della fine anticipata del blocco lungo il persorso di scansione (bibl. 4) consente di avere un numero frazionario di bit per simbolo e quindi di approssimare i requisiti richiesti dalla teoria della *Rate Distorsion.*

Con questo schema di codifica la quantizzazione uniforme risulta migliore di quella non uniforme. Supponendo inoltre i coefficienti DCT con uguale varianza, la quantizzazione scalare è quasi ottima, poiché richiede solo 0,254 bit per simbolo in più di quelli previsti dalla *Rate Distortion Theory* (bibl. 15). Il valore di 0,254 bit per simbolo corrisponde in pratica al possibile guadagno di «riem-

pimento dello spazio» previsto dalla teoria dell'«High Resolution» per il quantizzatore vettoriale. Nel caso pratico i coefficienti DCT non hanno tutti la stessa varianza e questo è sfruttato dalla codifica delle sequenze di zeri e della fine anticipata del blocco (bibl. 4); tale fatto comporta un'ulteriore riduzione del guadagno della quantizzazione vettoriale rispetto a quella scalare.

Risulta quindi discutibile l'opportunità di usare un quantizzatore vettoriale a meno che non vi sia la possibilità di sfruttare meglio l'eventuale presenza di memoria nella sorgente. Occorre rilevare che se le operazioni di quantizzazione e codifica sono separate, cioè il vocabolario da codificare non è direttamente quello fornito dal quantizzatore, è il codificatore a trarre vantaggio dalla memoria della sorgente.

6. Codifica vettoriale dei coefficienti DCT

Al fine di limitare il numero di parametri da considerare, lo studio e l'ottimizzazione sono stati fatti limitando l'algoritmo di codifica al solo modo intra-field.

L'ottimizzazione è stata effettuata utilizzando una sequenza costituita da pochi quadri che sono stati presi da diverse sequenze di prova. In tale modo si è ottenuto un campione sufficientemente rappresentativo delle sequenze da trattare.

L'indagine statistica e l'ottimizzazione sono state fatte a fattore di scalamento dei coefficienti DCT costante ed uguale a 60, 68, 76 corrispondenti ad una precisione massima dei coefficienti, escluso quello di continua, di circa 10, 9 e 8 bit (bibl. 30); in tale modo i bit rate medi complessivi, luminanza più crominanza, sono risultati di circa 30, 20, 10 Mbit/s, come si può vedere dalle tabelle riportate successivamente. I coefficienti DCT sono inoltre stati pesati mediante la matrice di visibilità definita dal CMTT/2 (bibl. 30).

6.1 STATISTICHE SULL'IMMAGINE

Una codifica entropica vettoriale dei coefficienti DCT è vantaggiosa rispetto alla codifica entropica scalare se la distribuzione del vocabolario risulta meno uniforme di quella del vocabolario scalare, cosa che avviene se esiste correlazione tra i simboli emessi dalla sorgente.

Il rapporto di correlazione r_{xy} tra due variabili x e y è dato da:

$$r_{xy} = E\{(x-\mu_x)\cdot(y-\mu_y)\}/(\sigma_x\sigma_y) \quad [12]$$

dove u_x, u_y, e σ_x, σ_y indicano, rispettivamente, il valore medio e la varianza delle due variabili x e y. Il rapporto di correlazione può variare tra -1 e $+1$; se r_{xy} vale 0 allora le due variabili sono scorrelate. Quando le variabili sono più di due si parla di matrice di correlazione i cui elementi, facilmente ricavabili mediante un'estensione della [12], sono tutti i possibili rapporti di correlazione tra le variabili; la matrice risulta simmetrica con diagonale unitaria.

Già si sapeva che i coefficienti DCT hanno valor medio nullo, tranne la continua, una distribuzione quasi Laplaciana bilatera con energia che dipende dalla loro posizione all'interno del blocco DCT (bibl. 27 e 22) e che i coefficienti DCT appartenenti allo stesso blocco di codifica sono tra loro pressoché scorrelati.

I risultati della statistica hanno confermato le considerazioni fatte precedentemente ed i rapporti di correlazione tra coefficienti anche ad energia alta (vicino al coefficiente di continua) sono risultati quasi nulli.

Rappresentando i singoli coefficienti DCT come prodotto di segno e modulo e facendo un'indagine statistica separata per ognuno di essi, si sono ottenuti i seguenti risultati: i segni sono a valor medio nullo ed hanno una correlazione praticamente nulla, i moduli dei coefficienti sono invece correlati tra loro e la correlazione è forte per i coefficienti vicini a quello di continua. I rapporti di correlazione più elevati ($0,4 \div 0,6$) si trovano vicini alla diagonale principale della matrice di correlazione e corrispondono a coefficienti adiacenti lungo il percorso di scansione del blocco. Gli altri valori della matrice di correlazione variano tra 0,2 e 0,5 e corrispondono ad elementi distanti tra loro lungo il percorso di scansione a zig-zag del blocco di codifica, anche se possono essere vicini all'interno del blocco stesso.

In base ai risultati di queste statistiche può esser conveniente un sistema di codifica che tratti separatamente segni e moduli in modo da sfruttare, per quanto possibile, la correlazione di questi ultimi.

Per non togliere simmetria alla distribuzione dei segni, il valore 0 è stato considerato come un valore di segno.

È stato inoltre verificato che la correlazione tra i moduli rimane elevata anche dopo l'eliminazione dei coefficienti nulli.

6.2 COSTITUZIONE DEI VETTORI E LORO ENTROPIA

Si descrivono ora i sistemi di codifica messi a confronto ed i risultati ottenuti sulla base della loro entropia per elemento d'immagine (pel).

Come riferimento è stato adottato l'attuale sistema utilizzante quindi la matrice di visibilità per la pesatura dei coefficienti DCT (bibl. 2) con i vocabolari, distinti per luminanza e crominanza, formati da tutti i possibili coefficienti, segno compreso, da tutte le possibili sequenze di zeri consecutivi lungo il percorso di scansione, «run length di zeri», e dal simbolo di fine blocco, «End Of Block»; non è stato considerato il coefficiente relativo alla componente continua.

I valori di entropia della componente di luminanza e di crominanza riportati in Tabella 1 si riferiscono a tre diversi fattori di scalamento.

Nel sistema di codifica vettoriale si è mantenuto lo stesso quantizzatore, la stessa matrice di visibilità e scansione del blocco DCT del sistema di riferimento. Analogamente a quanto viene fatto nel sistema preso come riferimento (bibl. 2), il blocco $8\cdot8$ viene trasformato in un vettore di 64 elementi mediante una scansione a zig-zag.

Il valore di ciascun elemento del vettore viene quindi diviso in due fattori moltiplicativi; il primo contenente il segno, includendo in tale parte anche il valore 0, ed il secondo il valore assoluto del coefficiente. Il vettore di 64 elementi contenente i segni viene suddiviso in vettori di dimensione più piccola; sono state considerate dimensioni varianti da 3 a 6.

TABELLA 1

ENTROPIA DEI COEFFICIENTI CODIFICATI SCALARMENTE E VETTORIALMENTE

	sistema riferim.			vettori segno e moduli scalari						vettori segno e vettori modulo					
	lum.	crom.	tot.	luminanza		crominanza		totale		luminanza		crominanza		totale	
scal	H/pel	H/pel	H/pel	H/pel	G.%	H/pel	G.%	H/pel	G%	H/pel	G.%	H/pel	G.%	H/pel	G.%
60	2,22	0,77	2,99	2,14	3,5	0,71	7,2	2,86	4,4	2,01	9,2	0,68	11,6	2,69	9,8
68	1,48	0,40	1,89	1,44	3,0	0,37	9,2	1,80	4,3	1,36	7,9	0,35	13,2	1,71	9.1
76	0,97	0,19	1,17	0,94	3,5	0,17	12,4	1,11	5,0	0,90	7,4	0,17	15,0	1,07	8,6

Nota: entropia delle componenti di luminanza e crominanza ed entropia totale dei coefficienti DCT quantizzati scalarmente e con codifica scalare dei coefficienti, con codifica vettoriale dei segni e scalare di coefficienti e con codifica vettoriale dei segni e dei coefficienti. Nella Tabella sono anche riportate le riduzioni di entropia dei sistemi vettoriali rispetto al sistema di codifica scalare preso come riferimento.

La continua viene considerata da sola, mentre il vettore più vicino ad essa ha sempre dimensione 3. In tale modo si ricopre esattamente il vettore di 64 elementi ottenuto dall'operazione di scansione del blocco. Al fine di utilizzare la diversa varianza dei coefficienti lungo il percorso di scansione, si suddividono i vettori in tre classi di cui la prima è costituita dal solo vettore di 3 elementi vicino alla componente continua. In questo modo si mantengono quasi completamente i vantaggi forniti dalla codifica della fine anticipata del blocco, EOB, che viene sempre inviato dopo l'ultimo vettore non nullo e della codifica delle sequenze di zeri.

In base ai risultati ottenuti dalle indagini statistiche si è rivelato vantaggioso introdurre una classificazione anche sull'energia complessiva del blocco DCT; da essa infatti si può dedurre tendenzialmente il numero di zeri e la posizione dell'EOB. In questa prima fase, i moduli dei coefficienti non vengono raggruppati in vettori, ma vengono codificati scalarmente.

L'entropia decresce molto lentamente all'aumentare della dimensione del vettore. Per motivi di uniformità con la dimensione dei vettori per la codifica dei moduli si sono scelti vettori di dimensione 4 con prestazioni praticamente uguali a quelle ottenibili con vettori segno di dimensione superiore. Poiché ogni elemento può assumere 3 valori, 1, − 1 e 0 si hanno in totale 81 differenti vettori di 4 elementi. In Tabella 1 si riportano, nelle apposite colonne, i valori dell'entropia per pel, relativi alle componenti di luminanza e crominanza ed il guadagno ottenuto rispetto al sistema di riferimento.

Al fine di sfruttare la correlazione esistente tra i moduli dei coefficienti DCT si è attuata la codifica vettoriale sia dei loro segni, includendo i coefficienti DCT nulli, descritta precedentemente, sia dei loro moduli. I coefficienti DCT nulli non vengono quindi più codificati come moduli per cui il numero di moduli da codificare vettorialmente non è fisso, ma varia da blocco a blocco. Alcuni coefficienti vicini alla continua, lungo il percorso di scansione, possono quindi essere codificati scalarmente al fine di poter riempire un numero intero di vettori con i restanti moduli.

Come esposto al paragrafo 2.2, il problema principale della codifica vettoriale esaustiva dei moduli risulta l'estensione del vocabolario che può essere molto grande e quindi di non pratica utilizzazione. Infatti, la precisione massima dei coefficienti DCT risulta di 11 bit escluso il segno per cui sarebbe necessario un vocabolario di 2^{22} simboli per una codifica esaustiva di vettori a sole 2 dimensioni. Si è quindi considerato una codifica vettoriale dei valori più bassi e quindi più probabili ed una codifica scalare per i valori elevati dei moduli.

Ciò avviene per mezzo di uno stadio scalare che opera nel modo illustrato in figura 11; quando un modulo del vettore in codifica raggiunge o supera il valore massimo (MAX) prefissato, la differenza viene codificata scalarmente e trasmessa al ricevitore senza che si renda necessaria alcuna informazione aggiuntiva. Tra tutti gli schemi praticamente realizzabili provati (2 stadi vettoriali, dimensione del vettore variabile, ecc.) questo ha fornito i risultati migliori. Al fine di non aumentare eccessivamente il numero di bit delle parole VLC associate ai vettori meno probabili si è posta la limitazione a 10000 simboli del vocabolario. Con tale restrizione i risultati migliori si sono ottenuti con vettori di dimensione 4; gli elementi del vettore possono assumere i valori da 1 a 10.

Come per la codifica dei vettori segno, anche per i vettori modulo è stata introdotta la classificazione energetica e quella di posizione.

I risultati in termini di entropia/pel del sistema complessivo «vettori segno e vettori modulo», entrambi di dimensione 4, e il suo guadagno percentuale rispetto al sistema di riferimento sono riportati in Tabella 1.
Dai valori di entropia si può vedere che la codifica vettoriale dei soli segni, incluso il valore zero dei coefficienti, riduce di circa il 3% l'entropia della componente di luminanza e di circa il 9% quella della componente di crominanza, e quindi ottiene una riduzione dell'entropia totale di circa il 4,5% rispetto al sistema di riferimento che codifica scalarmente i coefficienti DCT inclusi di segno e modulo. Si può notare che la variazione dell'entropia della crominanza agisce poco sulla variazione totale di entropia in quanto tale entropia è circa il 20% dell'entropia totale.

Codificando vettorialmente sia i segni che i moduli, con il requisito di un vocabolario, code book, non superiore a 10000 simboli, si ottiene un valore d'entropia inferiore a quella del sistema di riferimento, codifica scalare, di circa l'8% per la componente di luminanza e circa il 13% per la componente di crominanza. Tenendo conto del rapporto

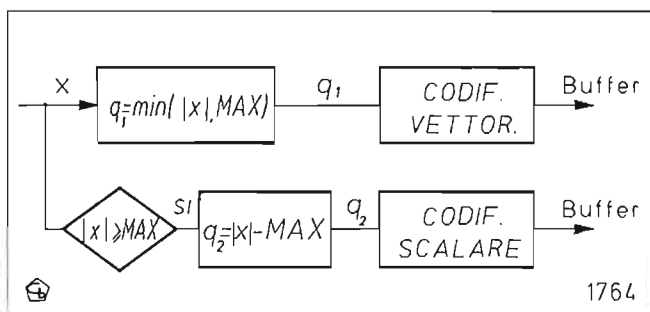



Fig. 11 — Schema di codifica a due stadi. Il primo stadio vettoriale codifica gli elementi aventi un valore del modulo inferiore a MAX, mentre il secondo stadio scalare codifica la differenza tra il modulo dei singoli elementi ed il valore MAX quando essa è positiva o nulla.

di entropia tra le componenti, si ottiene una riduzione di circa il 9% dell'entropia totale.

6.3 Schema di codifica vettoriale

Il sistema di codifica vettoriale complessivo può essere descritto per mezzo dello schema a blocchi riportato in figura 12.

In tale schema si possono distinguere le parti che hanno il compito di organizzare i coefficienti del blocco in vettori segno e vettori modulo, i quali saranno codificati da un VLC diverso per ogni vocabolario; la classificazione di posizione è implicitamente conosciuta anche dal ricevitore, mentre occorre trasmettere in parallelo l'informazione relativa alla classe di energia.

6.4 Risultati della codifica vettoriale dei coefficienti DCT

Il sistema di codifica è completato dall'assegnazione, ad ogni parola del vocabolario, di un simbolo di codice la cui lunghezza, in bit, dipende dalla probabilità della parola codificata.

I codici che possono essere adottati sono caratterizzati: dall'efficienza che possono raggiungere, cioè il rapporto tra l'entropia e il numero medio di bit, dalla robustezza contro gli errori residui di canale e dalla semplicità delle operazioni di co-decodifica (bibl. 4).

Nella valutazione del sistema di codifica vettoriale in termini di bit rate sono stati presi in considerazione due codici diversi: il B_2-Code ed il codice di Huffman. Il primo, già usato per il sistema di riferimento, è robusto contro gli errori di canale in quanto ne limita la propagazione, ha una semplice co-decodifica ed inoltre risulta di facile progettazione. Il codice di Huffman ha invece la migliore efficienza possibile, note le probabilità dei simboli del vocabolario, però è particolarmente sensibile agli errori ed inoltre le operazioni di co-decodifica risultano complesse; esso viene comunque considerato al fine di valutare il limite di efficienza della codifica vettoriale.

Le prove condotte con il sistema di riferimento hanno evidenziato un'efficienza ottima per il codice di Huffman e lievemente inferiore per il B_2-Code; inoltre entrambi i codici hanno dimostrato di non essere molto sensibili al variare della statistica delle immagini da codificare.

In Tabella 2 sono indicati i bit rate del sistema di riferimento ottenuti con il B_2-code ed il codice di Huffman e la loro efficienza.

Il sistema basato sulla codifica vettoriale ha fatto riscontrare una diminuzione di efficienza rispetto al sistema di riferimento. Tale riduzione è molto pronunciata nel caso in cui si utilizzi il B_2-code e questo è dovuto alla differenza tra la caratteristica della densità di distribuzione dei vettori da codificare e quella del B_2-code. La riduzione di efficienza del codice di Huffman nel caso vettoriale rispetto a quello scalare è dovuto alla maggior variabilità della distribuzione statistica dei vettori rispetto a quella dei singoli coefficienti DCT.

Fig. 12 — Schema a blocchi della codifica vettoriale dei segni e dei moduli dei coefficienti DCT. La codifica dei moduli è fatta a due stadi come indicato in figura 11.

TABELLA 2
CODIFICA SCALARE
(2a)

	B2-CODE					
	luminanza		crominanza		Totale	
scal.	bit/pel	ϵ %	bit/pel	ϵ %	bit/pel	ϵ %
60	2,36	93,9	0.81	95,1	3,17	94,2
68	1,57	94,7	0,43	92,8	2,00	94,3
76	1,02	95,1	0,23	86,3	2,25	93,5

(2b)

	HUFFMAN					
	luminanza		crominanza		Totale	
scal.	bit/pel	ϵ %	bit/pel	ϵ %	bit/pel	ϵ %
60	2,26	98,3	0.80	96,2	3,06	97,8
68	1,49	99,6	0,43	93,1	1,92	98,1
76	0,98	99,0	0,23	85,9	1,21	96,5

Nota: codifica scalare dei coefficienti DCT mediante B_2-Code, 2a, e mediante codice di Huffman, 2b. Nella tabella sono anche indicati i valori dell'efficienza di codifica.

In Tabella 3 sono riportati i valori di bit rate, efficienza e guadagno, rispetto alla codifica del sistema di riferimento, del sistema di codifica vettoriale dei segni e scalare dei moduli, mentre in Tabella 4 si riportano i corrispettivi valori del sistema di codifica vettoriale dei segni e moduli. Il numero di valori assunti da ogni elemento del vettore è stato ulteriormente ridotto da 10 a 6, ottenendo in tal modo 1296 differenti vettori, in quanto risultava eccessivamente lungo il tempo necessario a determinare il codice di Huffman per una tabella di 10000 elementi. I valori di efficienza riportati in Tabella 4 non si riferiscono quindi ai valori di entropia riportati in Tabella 1, ma a quelli, non riportati, relativi alla sorgente codificata con il codice di Huffman.

Dal confronto delle tabelle, si vede che il B_2-code non è adatto alla codifica vettoriale dei coefficienti DCT, in quanto la distribuzione statistica dei vettori è molto diversa dalla distribuzione delle parole del codice. Non è conveniente codificare vettorialmente i soli segni in quanto la riduzione di bit rate ottenuta risulta circa il 3%, a scapito di un notevole aumento di complessità.

Codificando vettorialmente sia i segni che i moduli si ottiene, utilizzando il codice di Huffman, una riduzione del bit rate rispetto al sistema con codifica scalare variante tra il 5 ed il 7% in dipendenza del fattore di trasmissione utilizzato. L'aumento di complessità di tale tipo di codifica e la sua minore robustezza contro gli errori residui di canale non sembrano controbilanciati dalla riduzione di bit rate ottenuta. Si deve comunque considerare che valori di efficienza prossimi a quelli riportati in tabella possono forse essere ottenuti mediante codici di tipo B, diversi dal B_2-code, ma con analogo comportamento nei confronti degli errori residui di canale.

TABELLA 3
CODIFICA VETTORIALE DEI SEGNI E SCALARE DEI MODULI

(3a)

	B-CODE								
	luminanza			crominanza			Totale		
scal.	b/pel	ε %	G.%	b/pel	ε %	G.%	b/pel	ε %	G.%
60	2,35	91,3	0.7	0,85	83,7	−5,3	3,20	89.3	−0,9
68	1,58	91,2	−0,7	0,46	78,8	−6,4	2,04	88,4	−2,0
76	1,06	88,7	−3,5	0,25	67,8	−11,5	1,31	84,7	−4,9

(3b)

	HUFFMAN								
	luminanza			crominanza			Totale		
scal.	b/pel	ε %	G.%	b/pel	ε %	G.%	b/pel	ε %	G.%
60	2,21	96,9	2,0	0,75	95,7	6,8	2,96	96,6	3,3
68	1,46	98,3	1,8	0,40	91,1	7,3	1,86	96,8	3,0
76	0,96	97,6	2,1	0,21	81,0	7,1	1,17	94,6	3,0

Nota: codifica vettoriale dei segni e scalare dei moduli dei coefficienti DCT mediante B_2-Code e mediante codice di Huffman. Nella tabella sono anche indicati i valori dell'efficienza di codifica e la riduzione di bit rate rispetto alla codifica scalare.

TABELLA 4
CODIFICA VETTORIALE DEI SEGNI E DEI MODULI

(4a)

	B2-CODE								
	luminanza			crominanza			Totale		
scal.	b/pel	ε %	G.%	b/pel	ε %	G.%	b/pel	ε %	G.%
60	2,46	81,0	−5,3	0,84	81,1	−3,6	3,32	81,0	−4,8
68	1,66	82,4	−5,8	0,45	78,0	−3,3	2,10	81,5	−5,2
76	1,10	82,1	−7,2	0,24	68,5	−7,1	1,34	79,7	−7,2

(4b)

	HUFFMAN								
	luminanza			crominanza			Totale		
scal.	b/pel	ε %	G.%	b/pel	ε %	G.%	b/pel	ε %	G.%
60	2,11	96,3	6,6	0,74	93,5	7,7	2,84	94,7	6,9
68	1,40	98,0	5,7	0,40	87,7	6,9	1,81	94,9	6,0
76	0,94	96,8	4,6	0,21	77,7	5,6	1,15	92,6	4,8

Nota: codifica vettoriale dei segni e dei moduli dei coefficienti DCT mediante B_2-Code e mediante codice di Huffman. Nella tabella sono anche indicati i valori dell'efficienza di codifica e la riduzione di bit rate rispetto alla codifica scalare.

7. Quantizzazione vettoriale dei coefficienti DCT

Si sono analizzate le prestazioni ottenute mediante quantizzazione vettoriale dei coefficienti DCT, interni ai singoli blocchi di codifica, abbinata alla codifica entropica dei vettori quantizzati in modo da consentire la regolazione fine del bit rate operabile con il buffer di trasmissione.

In questo contesto è stato scelto il quantizzatore vettoriale a reticolo utilizzante come cella base l'icosaedro troncato generalizzato alla dimensione 64.

Tale cella, che è ottima per le dimensioni 3 e 4 (bibl. 24), è infatti generalizzabile ad ogni dimensione basandosi sull'osservazione che le coordinate del baricentro sono o tutte pari o tutte dispari per ogni cella del reticolo.

Questa proprietà è alla base dell'algoritmo di quantizzazione, che risulta così oltremodo rapido e semplice. È infatti sufficiente effettuare due quantizzazioni ipercubiche del blocco, una per avere coordinate tutte pari l'altra per averle tutte dispari e scegliere tra le due possibili rappresentazioni quella che produce la distorsione minore.

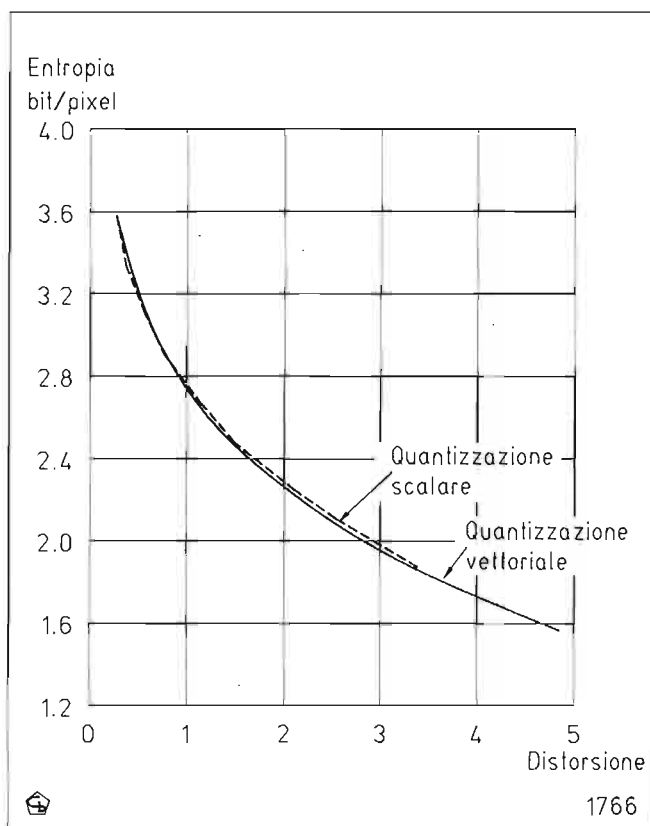
Come esposto precedentemente, il quantizzatore vettoriale con codifica entropica presenta un guadagno di riempimento dello spazio rispetto al quantizzatore scalare con codifica vettoriale entropica illustrato in figura 5b. Tale guadagno, inferiore a 0,254 bit per campione, si ha quando sono soddisfatte le ipotesi dell'«High Resolution Theory» ed in particolare quando i simboli emessi dalla sorgente sono scorrelati tra loro e con identica distribuzione statistica. Tale fatto comporta che la probabilità di una *n-pla* è data dal prodotto delle probabilità dei singoli elementi costituenti la *n-pla*. In particolare la probabilità di avere n simboli nulli consecutivi è uguale alla potenza n-esima della probabilità di un simbolo nullo. In questo

modo la codifica delle sequenze di zeri non risulta applicabile.

I coefficienti DCT interni al blocco di codifica sono scorrelati tra loro, ma non hanno una uguale distribuzione statistica. In particolare la probabilità di avere sequenze di n zeri è elevata e molto superiore alla potenza n-esima della probabilità di un solo zero. Tale fatto comporta una notevole efficienza della codifica di sequenze di zeri e di fine anticipata del blocco. Le celle di quantizzazione ottime secondo l'«High Resolution Theory», riducono la probabilità degli zeri in uscita dal quantizzatore e quindi riducono l'efficienza della codifica di sequenze di zeri.

Per ovviare a tale inconveniente i coefficienti DCT hanno subito una elaborazione prima della loro quantizzazione vettoriale. Tale elaborazione consiste nell'annullare i coefficienti con valore assoluto inferiore ad una determinata soglia in modo tale da non ridurre il numero degli zeri presenti all'uscita del corrispettivo quantizzatore scalare. Quinti il quantizzatore vettoriale a reticolo opera soltanto sui coefficienti non nulli del blocco. In tale modo il quantizzatore non è più rigorosamente uniforme e le sue prestazioni complessive risultano purtroppo circa uguali a quelle del quantizzatore scalare, come si può vedere in figura 13 dove sono riportate le curve distorsione/entropia dei due quantizzatori.

8. Conclusioni

Si sono analizzate le tecniche ed i vantaggi relativi alla codifica ed alla quantizzazione vettoriale dei coefficienti DCT.

Esistono due schemi fondamentali di codifica DCT, uno detto di bit allocation e l'altro detto di codifica entropica. Il primo codifica ciascun coefficiente con un numero di bit definito dalla «Rate Distorsion Theory» in base alla varianza del coefficiente ed alla distorsione accettata. Il quantizzatore utilizzato è non uniforme ed ottimo nel senso di Max e varia da coefficiente a coefficiente. Con tale tipo di schema la quantizzazione e codifica vettoriale presenta parecchi vantaggi, descritti nell'articolo, rispetto a quella scalare e tra questi la possibilità di ottenere, per ogni coefficiente, un'allocazione frazionaria del numero di bit.

Lo schema di codifica entropica utilizza un buffer di canale che retroagisce sulla caratteristica di quantizzazione. Tale metodo di codifica è stato descritto in vari articoli pubblicati su questa stessa rivista. Nel presente articolo si sono esaminate le possibilità di impiego su di esso della codifica e quantizzazione vettoriale.

La DCT con codifica entropica permette la codifica con un numero frazionario di bit per campione già nel caso scalare, grazie alla codifica di sequenze di zeri e di fine anticipata del blocco. La codifica vettoriale non presenta quindi i vantaggi legati alla «Rate Distorsion Theory».

La codifica vettoriale risulta vantaggiosa nel caso in cui i coefficienti DCT siano correlati. Purtroppo questo non si verifica se si considerano i coefficienti comprensivi di segno e modulo, ma si presenta invece tra i loro moduli. Si può quindi ottenere un guadagno di codifica scindendo i singoli coefficienti in modulo e segno che vengono poi separatamente codificati in modo vettoriale.

La riduzione di bit rate, 5 ÷ 7%, ottenuta utilizzando il codice di Huffman, non sembra sufficiente per passare ad un sistema di codifica notevolmente più complesso. Tale cambio non sembra conveniente neanche nel caso in cui, ampliando il vocabolario della codifica vettoriale, si raggiungesse una riduzione di bit rate del 9% che risulta pari alla riduzione di entropia. Inoltre il codice di Huffman risulta poco protetto contro gli errori residui di canale e quindi sarebbe conveniente individuare un B-code avente, con la codifica vettoriale, un'efficienza prossima a quella del codice di Huffman.

Inoltre, nello schema di DCT con codifica entropica, la quantizzazione vettoriale non presenta nessun vantaggio rispetto alla quantizzazione scalare.

Le ottimizzazioni sono state fatte con uno schema di codifica DCT che utilizza solo il modo intra-field. Si sono comunque fatte alcune verifiche, per quanto riguarda il quantizzatore a reticolo, con i coefficienti ottenuti dalla codifica di DCT ibrida ed i risultati non sono stati sostanzialmente diversi. Inoltre si sono fatte alcune prove di quantizzazione vettoriale formando i vettori con gli elementi coposizionati di blocchi differenti; i risultati ottenuti non sono stati migliori di quelli descritti.

(3957)

Fig. 13 — Entropia in funzione dell'errore di quantizzazione del segnale quantizzato in modo scalare e vettoriale con quantizzatore a reticolo.

BIBLIOGRAFIA

1. - DEL PERO R.: *Codifica del segnale televisivo numerico: metodi di riduzione della ridondanza.* «Elettronica e Telecomunicazioni», n. 1, 1989, pagg. 9-16.

2. - BARBERO M., STROPPIANA M.: *Codifica del segnale televisivo numerico: uso della trasformata coseno discreta.* «Elettronica e Telecomunicazione», n. 1, 1989, pagg. 17-29.

3. - BARBERO M., DEL PERO R., GIROMINI P.: *Codifica del segnale televisivo numerico: struttura di trama per un sistema basato su DCT.* «Elettronica e Telecomunicazioni», n. 1, 1990, pagg. 13-22.
4. - BARBERO M., BELLORA R., STROPPIANA M.: *Codifica del segnale televisivo numerico: codici a lunghezza variabile applicati ai coefficienti DCT.* «Elettronica e Telecomunicazioni», n. 1, 1990, pagg. 23-33.
5. - BARBERO M., CUCCHI S.: *Codifica del segnale televisivo numerico: architettura di un co-decodificatore HDTV utilizzante la DCT.* «Elettronica e Telecomunicazioni», n. 1, 1990, pagg. 3-11.
6. - *SPECIALE ITALIA '90.* «Elettronica e Telecomunicazioni», n. 3, 1990.
7. - BENEDETTO S., BIGLIERI E., CASTELLANI V.: *Digital Transmission Theory.* Prentice Hall, 1987.
8. - JOEL MAX: *Quantizing for Minimum Distorsion.* «IRE Transactions on Information Theory», vol. IT-6, March 1960.
9. - A. GERSHO: *Principles of Quantization.* «IEEE Transactions on Circuits and Systems», vol. CAS-25, No. 7, July 1978.
10. - T. BERGER: *Optimum Quantizers and Permutation Codes.* «IEEE Transactions on Information Theory», vol. IT-18, November 1972.
11. - ROGER C. WOOD: *On Optimum Quantization.* «IEEE Transactions on Information Theory», vol. IT-15, No. 2, March 1969.
12. - H. GISH, J. N. PIERCED: *Asymptotically Efficient Quantizing.* «IEEE Transactions on Information Theory», vol. IT-14, No. 5, September 1968.
13. - M. STROPPIANA: *Valore ottimo di restituzione del quantizzatore lineare e suo impiego nella codifica dei coefficienti DCT.* RAI - Centro Ricerche, Relazione Tecnica 88/4/I.
14. - J. MAKOUL, S. ROUCOS, H. GISH: *Vector Quantization in Speech Coding.* «Proceedings of the IEEE», vol. 73, No. 11, November 1985.
15. - T. D. LOOKABAUGH, R. M. GRAY: *High-Resolution Quantization Theory and the Vector Quantizer Advantage.* «IEEE Transactions on Information Theory», vol. IT-35, No. 5, September 1989.
16. - A. GERSHO: *On the Structure of Vector Quantizers.* «IEEE Transactions on Information Theory», vol. IT-28, No. 2, March 1982.
17. - P. A. CHOU, T. LOOKABAUGH, R. M. GRAY: *Entropy-Constrained Vector Quantization.* «IEEE Transations on Acoustics, Speech, and Signal Processing», vol. ASSP-37, No. 1, January 1989.
18. - Y. LINDE, A. BUZO, R. M. GRAY: *An Algorithm for Vector Quantizer Design.* «IEEE Transactions on Communications», vol. COM-28, No. 1, January 1980.
19. - R. M. GRAY: *Vector Quantization.* «IEEE Acoustics, Speech and Signal Processing Magazine», April 1984.
20. - T. R. FISCHER: *A Pyramid vector Quantizer.* «IEEE Transaction on Information Theory», vol. IT-32, No. 4, July 1986.
21. - T. R. FISCHER: *Geometric Source Coding and Vector Quantization.* «IEEE Transaction on Information Theory», vol. IT-35, No. 1, January 1989.
22. - R. C. REININGER, J. D. GIBSON: *Distributions of the Two Dimensional DCT Coefficients for Images.* «IEEE transactions on Communications», vol. COM-31, No. 6, June 1985.
23. - F. BELLIFEMINE, A. CHIMIENTI, R. PICCO: *Analysis of a 2d-dct coding scheme with motion compensation and vector quantization.* «Proceeding of SPIE», Vol. 1360, Lausanne, october 1990.
24. - A. GERSHO: *Asymptotically Optimal Block Quantization.* «IEEE transaction on Information Theory», vol. IT-25, No. 4, July 1979.
25. - CONWAY, SLOANE: *Voronoi Regions of Lattices, Second Moments of Polytopes, and Quantization.* «IEEE Transaction on Information Theory», vol. IT-28, No. 2, March 1982.
26. - DAVISSON: *Rate Distortion Theory and Applications.* «Proceedings of the IEEE», July 1972.
27. - BELLORA, DIMINO, MURATORI: *Hybrid DCT Comparison of the Statistics of DCT Coefficients and processing modes and with and without Motion Compensation.* «International Workshop on HDTV», Torino, Settembre 1989.
28. - W. CHEN, C. H. SMITH: *Adaptive Coding of Monochrome and Color Images.* «IEEE transactions on Communications», vol. COM-25, No. 11, Nov. 1977.
29. - K. AIZAWA, H. HARASCÖHIMA, H. MIYAKAWA: *Adaptive Vector Quantization of Picture Signals in Discrete Cosine Transform Domain.* «Electronics and Communications in Japan», Part 1, vol.70, No. 5, 1987.
30. - CMTT/2 - 172.

CODIFICA DEL SEGNALE TELEVISIVO NUMERICO: MASCHERAMENTO DEGLI ERRORI RESIDUI DI CANALE

M. STROPPIANA, N. ZENONI*

SOMMARIO — Il sistema di codifica DCT ibrido, sviluppato nell'ambito del progetto EU256, è stato illustrato in vari articoli precedenti. Il presente articolo analizza il comportamento di tale sistema in presenza di errori non riconosciuti dal FEC (Forward Error Correction) e indica la possibilità di migliorarne il comportamento in particolare con canali ad alto tasso di errore, condizione in cui il FEC non riesce più a lavorare efficacemente e quindi non è più in grado di garantire un'immagine non degradata. La distribuzione statistica dell'errore di canale può essere, random od a burst, mentre quella dell'errore residuo dopo il FEC è principalmente a burst, specialmente se si utilizza una decodifica di Viterbi. Si sono quindi considerate entrambe le distribuzioni statistiche d'errore e si è simulata una probabilità d'errore residuo, all'uscita del FEC, pari a 10^{-3}. Uno dei risultati più rilevanti, ottenuti sulla base di analisi statistiche e valutazioni soggettive, è stato il determinare che le principali informazioni sulla presenza degli errori si possono ricavare dalla struttura di trama, senza ulteriore aggiunta di bit di ridondanza.

*SUMMARY — **Digital TV signal coding: concealment of residual channel errors.** The hybrid DCT coding has been developed in the Eureka 256 framework and has been described in previous papers. This article analizes its behaviour with high rate channel errors, in which the FEC (Forward Error Correction) does not work properly and a high amount of errors is present at its output. The channel can be affected by random and burst errors, but only burst errors are present at the output of a FEC or a Viterbi decoder. Random and burst error distributions have been considered with a bit error rate equal to 10^{-3}. Hence, the conditions of presence and absence of the FEC has been taken into account. The investigation showed that the most annoying errors can be detected using the features of the VLC and of the video frame structure.*

1. Introduzione

L'algoritmo di codifica sviluppato in Eureka 256 utilizza la trasformata DCT ibrida. Ogni semiquadro dell'immagine viene suddiviso in blocchi 8×8 e poi, ad ognuno di tali blocchi (modo intra-field) o ad una loro differenza rispetto ad un blocco di predizione ottenuto dai semiquadri precedenti (modi inter-field ed inter-frame), si applica la trasformata coseno discreta (DCT). I valori trasformati sono quantizzati in modo adattativo in dipendenza del grado di riempimento di un buffer di canale, sono scanditi lungo un percorso a zig-zag del blocco e sono poi codificati mediante sequenza (run-length) di zeri e mediante codice a lunghezza variabile progettato in base alla distribuzione statistica dei coefficienti DCT quantizzati. La fine di ogni blocco è individuata da un simbolo di fine blocco (EOB) che ingloba anche l'ultima sequenza di zeri incontrata lungo il percorso di scansione (bibl. 1).

A causa della codifica a lunghezza variabile, l'eventuale errore di canale può propagarsi spazialmente; inoltre, esso può anche propagarsi temporalmente da quadro a quadro a causa della codifica ibrida, predizioni inter-field e inter-frame.

La propagazione temporale dell'errore può essere limitata forzando periodicamente il codificatore ad operare in modo intra-field eseguendo così un «refresh» dell'immagine. La propagazione spaziale dell'errore può invece essere ridotta sfruttando le caratteristiche della struttura di trama del sistema di codifica e le potenzialità del VLC impiegato. In particolare, utilizzando il B-code, l'errore può essere confinato all'interno del simbolo oppure propagarsi al solo simbolo successivo quando è colpito un bit di continuità del codice. Questo comporta, in particolare, la suddivisione di un simbolo di codice in altri due, oppure la fusione di due in uno.

A causa della suddivisione dell'immagine in blocchi 8×8 e dell'applicazione della DCT, l'errore su uno o più coefficienti DCT in un blocco si manifesta sull'immagine come una alterazione dell'intero blocco ed è quindi generalmente visibile.

Se l'errore invece genera o cancella un simbolo di fine blocco, si altera il numero di blocchi per strisciata e si genera una visibile traslazione dell'immagine lungo la strisciata (raggruppamento di 8 righe consecutive). Contando il numero di blocchi per strisciata si può individuare la presenza dell'errore, ma non la sua localizzazione all'interno della stessa. La localizzazione dell'errore può avvenire utilizzando due simboli di fine blocco generati da un generatore pseudo casuale azzerato all'inizio di ogni

* Dott. Mario Stroppiana e ing. Nicola Zenoni del Centro Ricerche RAI - Torino.
Dattiloscritto pervenuto alla Redazione il 28 aprile 1993.

quadro. Il ricevitore confronta la sequenza di fine blocco ricevuta con quella generata localmente ed individua la posizione del blocco cancellato o creato. In tale modo, si può ottenere una corretta sincronizzazione dell'intera strisciata. Inoltre il blocco mancante può essere sostituito con il blocco decodificato e corrispondente del quadro precedente (bibl. 2).

Le caratteristiche dell'algoritmo di codifica e della struttura di trama possono fornire informazioni utili per la rivelazione di ulteriori errori residui di canale. Infatti, l'errore può provocare simboli non ammessi dal VLC (B-code) usato, ad esempio parole più lunghe di 18 bit, generare valori di coefficienti DCT fuori dalla dinamica del quantizzatore, creare blocchi contenenti più di 64 coefficienti, ecc. Il presente articolo illustra proprio il modo di utilizzo di queste informazioni al fine di migliorare l'immagine affetta da errori. Una volta rivelato l'errore su una parte dell'immagine si pone il problema di come eliminarlo o mascherarlo. Una tecnica impiegata di mascheramento consiste nel sostituire la parte d'immagine errata con un'altra parte d'immagine simile, per esempio quella nella stessa posizione nel quadro precedente. Tale operazione è abbastanza delicata in quanto potrebbe degradare ulteriormente l'immagine invece di migliorarla. Infatti, in alcuni casi l'errore causa un'alterazione poco visibile dell'immagine, mentre la sua sostituzione potrebbe provocare un peggioramento percettibile. Risulta quindi importante individuare gli errori di trasmissione molto visibili, e su questi effettuare l'operazione di mascheramento; gli errori poco visibili vengono visualizzati.

L'attività di ricerca si è quindi svolta in più direzioni: si è determinato il peso degli errori sui singoli dati, coefficienti DCT, modi di codifica, fattore di scalamento dei coefficienti (detto nel seguito fattore di trasmissione), e si è migliorata la tecnica di mascheramento (miglioramento della predizione della parte d'immagine corrotta), sfruttando al massimo le informazioni fornite dalla struttura di trama e dal VLC per la rivelazione degli errori.

2. Sensibilità dei singoli dati all'errore di trasmissione

È stato condotto uno studio preliminare sul degradamento della qualità dell'immagine causato dalla presenza di errori sui coefficienti DCT, allo scopo di evidenziare due tipi diversi di dati: quelli che, colpiti da errore, producono difetti maggiormente visibili, ad esempio un errore che colpisce il coefficiente di continua, e quelli che producono difetti meno percettibili, ad esempio errori che colpiscono i coefficienti DCT relativi alle frequenze più elevate (bibl. 3 e 4).

Per questioni di comodità nel considerare il blocco DCT di un'immagine si farà riferimento alla matrice associata (**A**), di dimensioni 8×8; il generico coefficiente DCT della matrice sarà indicato con a_{ij}.

Sono stati introdotti due tipi distinti di errori:

- *errore sui coefficienti DCT dell'immagine, codificati con codice a lunghezza variabile.*
- *errore sulle informazioni aggiuntive: modi di codifica, fattori di trasmissione, che sono codificati con codici a lunghezza fissa* (bibl.2).

Sono state considerate separatamente le componenti di luminanza e crominanza.

2.1 Visibilità dell'errore sui singoli coefficienti DCT

Viene introdotto l'errore sul singolo bit di ogni coefficiente DCT individuando le seguenti situazioni di errore:

- $a_{ij}+1$: *errore introdotto sul bit meno significativo;*
- $a_{ij}+2$: *errore introdotto sul 1° bit a partire dal meno significativo;*
- $a_{ij}+4$: *errore introdotto sul 2° bit a partire dal meno significativo;*
- $a_{ij}+2^k$: *errore introdotto sul k-esimo bit a partire dal meno significativo.*

Nelle figure 1 e 2 è riportato per alcuni coefficienti (quelli ritenuti maggiormente rappresentativi) relativi alla componente di luminanza e di crominanza, il numero di bit più significativi che occorre proteggere da errore per garantire una qualità «2» o «3» dell'immagine secondo la scala di qualità CCIR a 5 gradini.

Per il coefficiente relativo alla componente continua si richiede la massima protezione; si suppone pertanto di non introdurre su questo alcun errore.

Fig. 1 — Matrice associata ai coefficienti DCT in un blocco di immagine relativi alla componente di luminanza; per alcuni di essi è indicato il numero di bit più significativi che occorre proteggere da errore. Il coefficiente a_{11}, relativo alla componente continua è indicato con la sigla DC (Direct Component) e si suppone di garantire per esso la protezione più accurata.

Fig. 2 — Matrice associata ai coefficienti DCT di un blocco di immagine relativi alla componente di crominanza; per alcuni di essi è indicato il numero di bit più significativi che occorre proteggere da errore. Il coefficiente a_{11}, relativo alla componente continua è indicato con la sigla DC (Direct Component) e si suppone di garantire per esso la protezione più accurata.

2.2 VISIBILITÀ DELL'ERRORE SULLE INFORMAZIONI AGGIUNTIVE

Il modo di codifica indica se un quadriblocco è codificato in intra-field, inter-field od inter-frame.

Sono state fatte alcune simulazioni per verificare l'effetto prodotto sull'immagine quando tale parametro risulta affetto da errore e, come era logico attendersi, si è verificato che tale informazione richiede il più alto grado di protezione.

Considerazioni analoghe valgono per il fattore di trasmissione, parametro che è legato all'occupazione del buffer e determina la finezza del quantizzatore assegnato a tutti i quadriblocchi appartenenti ad una striscia di immagine (gruppo di 8 righe video). Almeno i 6 bit più significativi, degli 8 impiegati per la rappresentazione, richiedono una protezione molto accurata.

2.3 IPOTESI SU UNA STRATEGIA ELEMENTARE DI MASCHERAMENTO

Sulla base dei risultati riportati precedentemente, relativi alla visibilità degli errori sui singoli coefficienti DCT (vedi figure 1 e 2), sono state realizzate alcune prove per valutare una semplice strategia di mascheramento, «concealment», degli errori che colpiscono coefficienti DCT differenti.

In caso di immagine affetta da errori si forzano a zero i coefficienti DCT relativi alle alte frequenze spaziali, lasciando inalterati i coefficienti appartenenti a sottomatrici triangolari di **A** che contengono la continua, coefficiente a_{11}, e quelli di bassa frequenza spaziale. È noto in particolare che l'occhio risulta meno sensibile alle frequenze spaziali più elevate e che le frequenze puramente orizzontali e verticali sono visivamente più importanti di quelle diagonali.

Si illustrano le prove effettuate indicando semplicemente l'ordine della sottomatrice di **A** i cui coefficienti non sono stati azzerati.

<u>*Luminanza*</u>

1) *sottomatrice triangolare di dimensione* 4×4
2) *sottomatrice triangolare di dimensione* 5×5
3) *sottomatrice triangolare di dimensione* 6×6

<u>*Crominanza*</u>

1) *sottomatrice triangolare di dimensione* 2×2
2) *azzeramento di tutti i coefficienti tranne* a_{11}

È stato verificato, su 11 frame della sequenza «Renata», che nel caso della luminanza è sufficiente mantenere 21 coefficienti su 64 diversi da 0 per garantire una qualità accettabile (condizione 3); mentre nel caso della crominanza è sufficiente mantenere 3 soli coefficienti diversi da 0 (condizione 1) (bibl. 3 e 4).

Coefficienti DCT che devono essere trasmessi sono ottenuti dalla matrice **A**, partendo dall'elemento di continua, secondo un particolare percorso di scansione a zig-zag (bibl. 5 e 6).

In realtà i risultati ottenuti in questa prima parte non forniscono utili suggerimenti per la trasmissione del segnale video a bassi bit-rate; ciò è attribuibile al fatto che a basso bit-rate, inferiore od uguale ad 1 bit/pel, molta ridondanza è già stata soppressa e pressoché tutta l'informazione coincide con quella più sensibile alla presenza di errori. Inoltre, bisognerebbe introdurre un codice capace di rilevare la presenza di errori in ogni blocco e questo richiederebbe l'inserimento di parecchi bit di ridondanza.

3. Sensibilità agli errori dei singoli bit del flusso di dati (data stream)

Si sono eseguite simulazioni su 11 frame della sequenza «Renata» (esattamente dal frame 2 al frame 12), ad un bit-rate di circa 1 bit/pel; per avere indicazioni attendibili sul comportamento del sistema in presenza di errori, questi sono stati introdotti su tutti i frame considerati eccetto il primo, che si suppone trasmesso correttamente.

Il sistema di codifica e la struttura di trama sono quelle definite in EU-256 (bibl. 1 e 2), il codice a lunghezza variabile utilizzato (VLC, Variable Length Code) è il B-code (bibl. 7).

Le simulazioni sono state realizzate a fattore di trasmissione costante e si sono articolate in due fasi: senza compensazione e con compensazione del movimento.

In entrambi i casi gli errori sono stati introdotti con una probabilità di occorrenza di 10^{-3} e distribuiti in modo casuale (errori random) ed a pacchetti (errori a burst). La durata dei burst di errore è risultata, mediamente, inferiore a 20.

Gli errori introdotti possono produrre variazioni dei valori dei coefficienti DCT quantizzati e codificati oppure una variazione del loro numero.

Il B-code è un codice a lunghezza variabile ed è caratterizzato da una struttura particolarmente semplice: le parole di codice sono costituite da coppie di bit e la loro lunghezza varia da 2 a 18. Il primo bit di ogni coppia è detto bit di continuità ed indica se la parola di codice continua o finisce, mentre il secondo bit è un bit di informazione.

Anche le sequenze di zeri sono codificate con parole appartenenti al B-code.

In particolare, i coefficienti che si trovano alla fine del percorso di scansione del blocco si riferiscono alle componenti di alta frequenza e quindi hanno una probabilità elevata di essere nulli. Si preferisce quindi non inviare la parola corrispondente all'ultima sequenza di zeri, se presente, ma utilizzare un simbolo opportuno per indicare la fine del blocco; tale simbolo di fine blocco è indicato con il simbolo EOB (End Of Block).

In relazione al tipo particolare di codice utilizzato è risultato immediato discriminare tra due situazioni di errori a seconda che il bit affetto da errore fosse un bit di continuità o di informazione.

Nel primo caso la parola errata può dar luogo a due parole di codice, oppure due parole di codice possono essere trasformate in un'altra parola di codice; nel secondo caso, invece, l'errore non si propaga, ovvero rimane invariata la corretta divisione delle parole, ma viene decodificato un simbolo diverso da quello originario.

Il sistema di co-decodifica esaminato è in grado di rilevare, in fase di decodifica, alcune situazioni di errore mediante appositi controlli resi possibili dalle proprietà del VLC, del quantizzatore e delle parole di EOB utiliz-

zate; in questi casi viene forzato automaticamente il mascheramento del blocco errato mediante sostituzione con il corrispondente relativo al field o frame precedente. Vengono classificati differenti tipi di errore in dipendenza del modo con cui essi vengono rivelati.

L'errore può generare la perdita o la creazione di EOB (errori di tipo 1 e di tipo 2), può dar luogo alla generazione di un blocchetto che contiene più di 64 coefficienti (errore di tipo 3), può generare parole di codice non presenti nella tabella di codifica (errore di tipo 4).

Inoltre il quantizzatore impiegato presenta le seguenti caratteristiche:

1) *il campo dei valori possibili varia tra -639 e 639, mentre il valore della parole di codice disponibili varia da -733 a 733 [5,6];*
2) *il massimo valore dei coefficienti DCT, dopo lo scalamento inverso, è espresso su 12 bit.*

Se le condizioni *1* e *2* non sono verificate, il blocchetto in esame viene inoltre segnalato come errato (errori di tipo 5 e tipo 6).

Tutte queste situazioni di errore sono automaticamente riconosciute dal decodificatore e danno luogo ad una segnalazione di blocco errato.

Può verificarsi che in un blocchetto siano presenti errori multipli; in tal caso viene prodotta una sola segnalazione di errore; ciò spiega la leggera discrepanza che talvolta si osserva confrontando tra loro i numeri riportati nella prima colonna delle tabelle.

3.1 SISTEMA DI CODIFICA SENZA COMPENSAZIONE DI MOVIMENTO

Le simulazioni sono state condotte a fattore di trasmissione costante uguale a 80, corrispondente ad un bit-rate medio di circa 0,9 bit/pel, ed i risultati sono relativi ad errori distribuiti sia in modo random che a burst.

I bit delle singole parole del codice B-code si possono differenziare in due classi: i bit di continuità e quelli d'informazione. Come già detto nell'introduzione e come riportato in bibliografia 2, l'effetto degli errori si può manifestare in modo differente a seconda della classe di appartenenza del bit colpito. Si è quindi pensato di analizzare distintamente gli errori che colpiscono i bit di continuità, da quelli che colpiscono i bit d'informazione.

Per ogni ciclo di simulazioni si sono pertanto analizzate le seguenti 3 configurazioni di errore:

TABELLA 1

STATISTICHE RELATIVE A 3 CONFIGURAZIONI DISTINTE DI ERRORI RANDOM: SUI SOLI BIT DI CONTINUITÀ, SUI SOLI BIT DI INFORMAZIONE E SU ENTRAMBI I TIPI DI BIT

Tipo di bit errato	blocchi errati introdotti	blocchi errati rivelati	blocchi errati non rivelati	blocchi errati rivelati (%)
continuità	1855	656	1199	35
informazione	1838	486	1352	26
entrambi	3645	1140	2505	31

TABELLA 2

STATISTICHE RELATIVE A 3 CONFIGURAZIONI DISTINTE DI ERRORI A BURST: SUI SOLI BIT DI CONTINUITÀ, SUI SOLI BIT DI INFORMAZIONE E SU ENTRAMBI I TIPI DI BIT

Tipo di bit errato	blocchi errati introdotti	blocchi errati rivelati	blocchi errati non rivelati	blocchi errati rivelati (%)
continuità	892	447	445	50
informazione	933	415	518	44
entrambi	1066	544	522	51

- *errori sui soli bit di continuità ($P(e) = 0,5\ 10^{-3}$)*
- *errori sui soli bit di informazione ($P(e) = 0,5\ 10^{-3}$)*
- *errori su entrambi i tipi di bit ($P(e) = 10^{-3}$)*

Nelle tabelle 1 e 2 sono riportati il numero totale di blocchi affetti da errore, il numero di blocchi in cui si è rivelato l'errore ed il numero di blocchi in cui l'errore non è stato rivelato, rispettivamente nelle tre condizioni di errore indicate precedentemente. La distribuzione degli errori è stata considerata random in tabella 1 ed a burst in tabella 2.

Dalla tabella 1 risulta che la percentuale di blocchi errati rivelati è pari a circa il 31%; inoltre se gli errori colpiscono i bit di continuità viene rivelata una percentuale di blocchi errati (circa 35%) sensibilmente maggiore che nel caso in cui siano colpiti solo i bit di informazione (circa 26%). Questi risultati non sono tali da giustificare la scelta di un differente livello di protezione per i due tipi di bit. Valutazioni soggettive hanno invece evidenziato che gli effetti prodotti dagli errori non rivelati sui bit di continuità sono più fastidiosi di quelli prodotti da errori sui bit di informazione.

Nella tabella 2 aumentano decisamente le percentuali dei blocchi errati rivelati. Ciò accade, essenzialmente perché la distribuzione a burst di errori tende a concentrare un numero elevato di errori su uno stesso blocco, riducendo sensibilmente il numero di blocchi errati introdotti. Anche in questo caso si rivela una percentuale sensibilmente superiore di blocchi affetti da errore quando l'errore colpisce il bit di continuità; d'altra parte, il degradamento provocato dai blocchi in cui non è stato rivelato l'errore è, in quest'ultimo caso, maggiore del degradamento provocato dai blocchi con errore sui bit d'informazione.

Nelle tabelle 3 e 4 si riporta la frequenza dei vari tipi

TABELLA 3

FREQUENZA STATISTICA DEI TIPI D'ERRORE RANDOM RIVELATI DAL DECODIFICATORE. ERRORI SUI SOLI BIT DI CONTINUITÀ, SUI SOLI BIT DI INFORMAZIONE E SU ENTRAMBI I TIPI DI BIT

Tipo di bit errato	tipo 1	tipo 2	tipo 3	tipo 4	tipo 5	tipo 6	errori rivelati
continuità	484	125	28	1	3	15	656
informaz.	367	89	1	0	0	29	487
entrambi	851	213	29	1	3	43	1140

TABELLA 4
FREQUENZA STATISTICA DEI TIPI D'ERRORE A BURST RIVELATI DAL DECODIFICATORE: ERRORI SUI SOLI BIT DI CONTINUITÀ, SUI SOLI BIT DI INFORMAZIONE E SU ENTRAMBI I TIPI DI BIT

Tipo di bit errato	tipo 1	tipo 2	tipo 3	tipo 4	tipo 5	tipo 6	errori rivelati
continuità	363	57	12	0	2	13	447
informaz.	324	67	6	0	0	18	415
entrambi	452	68	10	0	1	13	544

d'errore che hanno originato un blocco errato rivelato dal decodificatore; la tabella 3 si riferisce ad errori random, e la tabella 4 ad errori a burst. Il tipo d'errore è già stato parzialmente descritto al paragrafo 3; una classificazione completa è riportata in tabella 5.

TABELLA 5
DESCRIZIONE DEI TIPI DI ERRORE

	Legenda
tipo 1	Perdita di simboli di fine blocco (EOB)
tipo 2	Creazione di nuovi simboli di fine blocco (EOB)
tipo 3	Blocchetto con più di 64 coefficienti
tipo 4	Parola di codice non presente in tabella
tipo 5	Valore coeff. DCT quantizzato non permesso
tipo 6	Valore coeff. DCT superiore range massimo
tipo 7	Errore su coeff. DCT relativo a comp. continua
tipo 8	Errore su bit di continuità
tipo 9	Errore su bit di informazione

Molto rilevanti dal punto di vista del deterioramento della qualità soggettiva dell'immagine sono gli errori che colpiscono il coefficiente DCT relativo alla componente continua; tale genere di errore produce infatti blocchetti uniformi.

In tabella 6 è riportato il numero di blocchetti affetti da tale errore distribuito, in modo casuale, solo sui bit di continuità, solo su quelli d'informazione oppure su entrambi.

TABELLA 6
STATISTICHE RELATIVE A 3 CONFIGURAZIONI DISTINTE DI ERRORI RANDOM SULLA CONTINUA: SUI SOLI BIT DI CONTINUITÀ, SUI SOLI BIT DI INFORMAZIONE E SU ENTRAMBI I TIPI DI BIT

Tipo di bit errato	totale errori introdotti	errore tipo 7	errore tipo 7 (%)
continuità	1855	268	14,5
informazione	1838	283	15,5
entrambi	3645	546	15

Dopo aver analizzato i vari tipi di errore, che a volte possono deteriorare anche pesantemente la qualità dell'immagine, si sono considerate tecniche di mascheramento dell'errore.

Tali tecniche consistono nel sostituire un blocco od un gruppo di blocchi affetti da errori con una predizione opportuna, riducendo il degradamento visibile dovuto agli errori residui. I valori di predizione sono comunque disponibili nel decodificatore; infatti durante il normale funzionamento del codificatore si utilizzano, come predizione, i campioni video dei due semiquadri precedenti (modi inter-field e inter-frame). I campioni ottenibili da dati affetti da errore possono essere sostituiti dai campioni video corrispondenti appartenenti al quadro precedente se il blocco errato è stato codificato in modo inter-frame, oppure da quelli appartenenti al semiquadro precedente se il blocco errato è stato codificato in modo intra-field o inter-field.

In questa fase si è eseguita l'operazione di mascheramento a livello di singolo blocco nei casi in cui il decodificatore è in grado di riconoscere l'errore in modo automatico (per gli errori classificati come tipo 1, tipo 2, ..., tipo 6); successivamente è stato forzato il mascheramento, a livello di blocco, anche nel caso di errori di tipo 7 e per errori che colpiscono i bit di continuità. La qualità dell'immagine ottenuta, applicando congiuntamente i vari tipi di mascheramento, produce l'annullamento delle principali fonti di disturbo e risulta notevolmente migliorata, come si può vedere dalle fotografie riportate in Appendice A, anche se non può essere giudicata ancora di livello soddisfacente.

È evidente che, in presenza di movimento, la sostituzione operata comporta comunque un degradamento, che si localizza principalmente nelle aree dell'immagine caratterizzate da elevato dettaglio. Infatti, a differenza del caso di decodifica secondo i normali modi inter in cui si utilizza l'errore di predizione trasmesso, in questo caso si sostituiscono semplicemente i campioni video ritenuti errati con i valori di predizione, generalmente diversi da quelli effettivi.

3.2 SISTEMA DI CODIFICA CON COMPENSAZIONE DI MOVIMENTO

In questa fase si valuta il miglioramento dell'immagine affetta da errori, in seguito all'applicazione delle tecniche di mascheramento effettuate tenendo conto dei vettori movimento valutati per i singoli quadriblocchi.

I vettori movimento utilizzati per l'operazione di mascheramento degli errori sono gli stessi utilizzati nel sistema di codifica per il modo inter-frame. Il blocco affetto da errore viene sostituito da quello preso dal frame precedente lungo la direzione indicata del vettore movimento, se il blocco è stato codificato nel modo inter-frame, oppure viene sostituito con quello interpolato del field precedente se il blocco è stato codificato in modo intra-field o inter-field.

Dalle fotografie riportate in Appendice B si può vedere il miglioramento del processo di mascheramento ottenibile utilizzando i vettori movimento.

L'applicazione della compensazione del movimento nel

sistema di codifica aumenta notevolmente l'efficienza di codifica, riducendo il bit-rate medio a parità di fattore di trasmissione e quindi a parità di rapporto S/N sull'immagine codificata. Si sono quindi eseguite varie simulazioni a differenti fattori di trasmissione costante; i valori corrispondenti di bit-rate medio sono riportati in tabella 7.

TABELLA 7
FATTORI DI TRASMISSIONE UTILIZZATI NELLE SIMULAZIONI E BIT-RATE ASSOCIATI

FT	bit rate
68	1,3
72	1
80	0,6

È evidente che per poter fare un confronto a parità di bit-rate con i risultati ottenuti senza compensazione del movimento, bisogna fare riferimento alla simulazione ottenuta nel caso FT = 72.

Le simulazioni analizzate in questa sezione sono state effettuate a fattore di scalamento, e quindi di trasmissione, costante e con la tecnica di compensazione del movimento, che utilizza per determinare i vettori movimento un'area di ricerca (Search Area) pari a +/- 15,5 pixel nella direzione orizzontale e +/- 7,5 righe nella direzione verticale. Si sono ancora presi come riferimento 11 frame della sequenza «Renata», che verranno utilizzati anche nel seguito.

Distingueremo anche qui due casi per la statistica degli errori introdotti: errori distribuiti in modo casuale ed a pacchetti con durata dei pacchetti di errore mediamente inferiore a 20. Entrambe le distribuzioni di errore hanno una probabilità pari a 10^{-3}.

In tabella 8 si riportano le statistiche relative a tutti i tipi di errori introdotti per i diversi valori di FT sopra indicati. Verranno indicati con tipo 8 e tipo 9 gli errori che colpiscono rispettivamente i bit di continuità e di informazione e che non originano tipi d'errore classificati da 1 a 7 (vedere tabella 5).

Il simulatore sviluppato per la parte di sistema relativa all'inserimento degli errori, fa corrispondere dunque, in modo univoco, ad ogni blocco errato una segnalazione appropriata; nel caso in cui un errore sia riconosciuto contemporaneamente da due diversi segnalatori di tipo, si risolve l'ambiguità segnalando l'errore di tipo inferiore, ad esempio un errore che cade sul coefficiente di continua viene indicato come tipo 7, indipendentemente dal fatto che sia interessato da errore un bit di continuità o di informazione. Inoltre può verificarsi che all'interno di un blocco cadano errori multipli; anche in tal caso viene prodotto, per quel blocco, la sola segnalazione di errore relativa al tipo inferiore.

Nella tabella 8 si riporta la distribuzione dei blocchi errati in funzione del tipo d'errore verificatosi, facendo riferimento ad una statistica di errore di tipo random. La tabella 9 riporta la statistica dei tipi d'errore nei vari blocchi con errori distribuiti a burst.

I risultati riportati nelle tabelle 7, 8 e 9 indicano una diminuzione del numero di blocchi affetti da errore al diminuire del bit-rate. La percentuale di errori rivelati aumenta al diminuire del bit-rate. In particolare aumenta

TABELLA 8
STATISTICHE RELATIVE AI VARI TIPI DI ERRORE RANDOM, CON PROBABILITÀ P(E) = 10^{-3} SU 10 FRAME DELLA SEQUENZA «RENATA», CON DIFFERENTI VALORI DEL FATTORE DI TRASMISSIONE

	FT = 68		FT = 72		FT = 80	
tipo di errore	n. blocchi errati	% blocchi errati	n. blocchi errati	% blocchi errati	n. blocchi errati	% blocchi errati
tipo 1	940	19	922	24	836	34,5
tipo 2	339	6,8	245	6,5	123	5
tipo 3	55	1	21	0,5	6	—
tipo 4	0	—	0	—	0	—
tipo 5	1	—	1	—	3	—
tipo 6	5	—	4	—	21	1
tipo 7	436	8,7	394	10,2	430	18
tipo 8	1541	31	1082	28	452	19
tipo 9	1681	33,5	1188	30,8	547	22,5
Tot. blocchi errati	4998	100	3875	100	2418	100

TABELLA 9
STATISTICHE RELATIVE AI VARI TIPI DI ERRORE A BURST, CON PROBABILITÀ P(E) = 10^{-3} SU 10 FRAME DELLA SEQUENZA «RENATA», CON DIFFERENTI VALORI DEL FATTORE DI TRASMISSIONE

	FT = 68		FT = 72		FT = 80	
tipo di errore	n. blocchi errati	% blocchi errati	n. blocchi errati	% blocchi errati	n. blocchi errati	% blocchi errati
tipo 1	419	30,5	465	40	418	53,5
tipo 2	113	8,5	54	4,5	42	5,5
tipo 3	31	2	12	1	6	0,5
tipo 4	0	—	0	—	0	—
tipo 5	1	—	1	—	1	—
tipo 6	3	—	6	0,5	5	0,5
tipo 7	183	13,5	193	16,5	141	18
tipo 8	538	39,5	377	32,5	142	18
tipo 9	83	6	57	5	29	4
Tot. blocchi errati	1371	100	1165	100	784	100

il numero di errori rivelati come tipo 1, perdita di EOB, e tipo 7, errore sulla continua. Questo è dovuto al fatto che il simbolo di fine blocco ed il valore della continua hanno un numero di bit pressoché indipendente dal valore del bit-rate e quindi la percentuale di bit ad essi destinata aumenta con la riduzione del bit-rate.

Si può infine notare che la percentuale di errori rivelati è decisamente più elevata nel caso di errori a burst, che risultano quindi meno critici di quelli random.

4. Mascheramento dell'errore residuo

Si è cercato di migliorare la qualità dell'immagine affetta da errore introducendo l'operazione di mascheramento, utilizzante i vettori movimento, a porzioni definite di immagine.

A parte il caso di errori riconosciuti dal decodificatore, per i quali scatta la procedura di mascheramento in modo automatico che può quindi essere attuata a livello di blocco o di quadriblocco, occorre applicare il mascheramento anche agli errori di tipo 7, che risultano particolarmente fastidiosi, e, per quanto possibile, anche a quelli di tipo 8.

Il vero problema consiste però nel fatto che per rilevare gli errori di tipo 7 e di tipo 8 bisogna prevedere un'aggiunta di ridondanza con conseguente riduzione dell'efficienza di codifica.

Per ridurre al minimo il prezzo da pagare in termini di bit-rate si è considerato il quadro televisivo suddiviso non già in blocchi o quadriblocchi, ma in porzioni più ampie, contenenti un multiplo intero di blocchetti elementari, detti macroblocchi e si è applicato su questi varie tecniche di mascheramento. Per gli errori di tipo 1...6 si è ritenuto migliore operare il mascheramento a livello di quadriblocco anziché di blocco. Infatti si sono ottenuti miglioramenti soggettivi, giustificati dal fatto che la sostituzione a livello di blocco separa le informazioni di luminanza e crominanza relative alla medesima porzione di immagine e rende le stesse non coerenti tra loro. In particolare si sono considerati due tipi di macroblocchi, costituiti rispettivamente da gruppi di quattro blocchi, detti quadriblocchi, e dai gruppi di 8 quadriblocchi.

4.1 ANALISI DELLE DIFFERENTI STRATEGIE

Le strategie di mascheramento operano nel seguente modo:

- *se il quadriblocco è processato in modo intra-field o inter-field, esso è sostituito con i valori di predizione ottenuti dai campioni video appartenenti al semiquadro immediatamente precedente;*
- *se il quadriblocco è processato in modo inter-frame con compensazione del movimento, esso è sostituito con i valori di predizione ottenuti dal blocco corrispondente lungo la direzione del movimento e appartenente al quadro precedente.*

In entrambi i casi si deve assegnare ad ogni macroblocco una certa quantità di bit di informazione aggiuntivi per rivelare il tipo di errore che si vuole mascherare. Per esempio, per il rivelamento di errori di tipo 7, errore sul valore della continua, riteniamo siano sufficienti 5 bit di protezione per quadriblocco corrispondenti a meno di 0,04 bit/pel. In tale caso la probabilità teorica di non rivelazione dell'errore è circa uguale a 3×10^{-2}.

Si sono quindi confrontati i seguenti tipi di mascheramento:

1) *a livello di quadriblocco per gli errori rilevati dal decodificatore (tipo 1, . , tipo 6);*
2) *a livello di quadriblocco per ogni errore dal tipo 1 al tipo 6 ed errori di tipo 7;*
3) *a livello di quadriblocco per ogni errore dal tipo 1 al tipo 6 ed errori di tipo 8;*
4) *a livello di quadriblocco per ogni errore dal tipo 1 al tipo 8;*
5) *a livello di quadriblocco per gli errori dal tipo 1 al tipo 6 ed a livello di 8 quadriblocchi per gli errori di tipo 7 e di tipo 8.*

Le simulazioni hanno rilevato che il numero di quadriblocchi per frame errati diminuisce all'aumentare del fattore di trasmissione e quindi al ridursi del bit-rate, come si può dedurre dalle tabelle 8 e 9 e che tale numero è decisamente inferiore con la distribuzione a burst degli errori. Ad un bit-rate di circa 1 bit/pel, fattore di trasmissione 72, la percentuale di blocchi errati in cui è stato rivelato l'errore e quindi è stato eseguito il mascheramento varia dal 31% (mascheramento indicato al punto 1) al 70% (mascheramento indicato al punto 4). Nelle stesse condizioni, ma con errore a burst le percentuali variano, rispettivamente, dal 46% al 95%.

Nel caso di mascheramento eseguito su 8 quadriblocchi non si è definita la percentuale di 8 quadriblocchi in cui si è rilevato l'errore. L'area su cui si esegue l'operazione di mascheramento risulta però estesa a gran parte dell'immagine, compromettendo così notevolmente l'efficienza della tecnica di «refresh». La qualità globale dell'immagine risulta perciò insoddisfacente, come esposto nel paragrafo successivo.

4.2 VALUTAZIONI SOGGETTIVE

Le considerazioni sotto riportate sono valide per i 2 tipi di configurazione di errore analizzati; la qualità soggettiva risulta migliore, a parità di condizioni, nel caso di errore a burst.

- *Come si può vedere nell'appendice 3, la strategia di mascheramento di 8 quadriblocchi presenta seri problemi sui contorni di oggetti in movimento e quindi non migliora la qualità dell'immagine.*
- *Nel caso della strategia indicata al punto 4 del paragrafo 4.1, la situazione migliora decisamente rispetto al caso precedente. Le aree dell'immagine caratterizzate da lento movimento non sono degradate, mentre i contorni di oggetti in rapido movimento danno luogo in fase di ricostruzione ad alcuni difetti, limitati però a piccole porzioni di immagine (dimensione del quadriblocco), la cui rilevanza visiva è ritenuta poco disturbante.*
- *Con la strategia indicata al punto 1, rilevazione dei tipi 1-6 d'errore, la qualità dell'immagine risulta piuttosto buona per i dettagli in movimento, mentre la presenza di errori sui coefficienti di continua disturba apprezzabilmente la visibilità delle zone di immagine a sfondo*

uniforme o quasi uniforme. Tale risultato è estremamente significativo e dimostra che senza alcuna aggiunta di ridondanza, ma sfruttando le sole peculiarità del decodificatore e le informazioni associate ai vettori movimento si può ottenere una qualità quasi accettabile dell'immagine anche in presenza di alti tassi di errore.

- *Si ottiene un notevole miglioramento, in termini di qualità soggettiva, con la soluzione indicata al punto 2 (rivelazione errore sulla continua).*

Le valutazioni riportate sulla qualità soggettiva ottenibile con le varie strategie di mascheramento sono verificabili esaminando le fotografie riportate in Appendice C.

5. Conclusioni

Lo studio condotto sul sistema di co-decodifica sviluppato nell'ambito del progetto EU256, a bassi bit-rate (circa 1 bit/pel) ed ipotizzando un alto tasso di errore sul canale ($P(e) = 10^{-3}$), ha contribuito ad ottenere una serie di utili indicazioni in materia di statistica degli errori e del loro mascheramento.

Innanzitutto si è potuto valutare in che misura ed in quale modo intervengono sull'immagine gli errori che colpiscono i bit di continuità e di informazione, concludendo che ai primi sono collegati i maggiori effetti di degradamento della qualità.

Particolare risalto hanno poi gli errori che colpiscono, all'interno di ogni blocco DCT, i coefficienti relativi alla componente continua.

Queste indicazioni hanno suggerito lo studio di differenti strategie di mascheramento, che ha portato a selezionare la tecnica, indicata al punto 2 del paragrafo 4, come il migliore compromesso tra la qualità soggettiva finale ottenibile e la quantità di bit rate aggiuntiva necessaria per il rilevamento degli errori.

È risultato comunque, senza dubbio, preferibile il tipo di mascheramento esteso a livello di singolo quadriblocco anziché a 8 quadriblocchi.

In definitiva si ritiene, sulla base dei dati raccolti, che è conveniente attuare una strategia di mascheramento a livello di quadriblocco per ogni errore individuato dal decodificatore (tipo 1, ..., tipo 6) e per errori di tipo 7: si è visto infatti che la qualità ottenibile in questo modo si può considerare accettabile e ciò a fronte di un incremento di ridondanza non rilevante.

Se si estende la tecnica di mascheramento sui singoli quadriblocchi anche agli errori di tipo 8 e di tipo 9 si ottiene un miglioramento della qualità soggettiva, ma tale scelta richiede un aumento elevato di ridondanza.

D'altra parte l'idea di applicare il mascheramento a gruppi di 8 quadriblocchi ha messo in luce forti problemi localizzati principalmente sulle parti dell'immagine in movimento.

I risultati ottenuti in questo lavoro, relativi a soli 10 frame della sequenza «Renata», si ritengono sufficientemente significativi e le conclusioni ottenute sono considerate ragionevolmente estendibili anche al caso di altre sequenze.

(4088)

APPENDICE A

Le fotografie riportate sono relative al frame 12 della sequenza «Renata», senza compensazione del movimento e con FT = 80 corrispondente ad un bit-rate medio di 0,9 bit/pel, in presenza di errore con distribuzione uniforme (errore random), con probabilità $P(e) = 10^{-3}$ e con differenti strategie di mascheramento degli errori.

Il mascheramento degli errori è stato ottenuto sostituendo il blocco errato con quello coposizionato appartenente al field od al frame precedente.

Fig. 1A — Nessun mascheramento degli errori.

Fig. 2A — Mascheramento degli errori di tipo da 1 a 6 (vedere tabella 5).

Fig. 3A — Mascheramento degli errori di tipo da 1 a 7 (vedere tabella 5).

Fig. 4A — Mascheramento degli errori di tipo da 1 a 8 (vedere tabella 5).

APPENDICE B

Le fotografie riportate sono relative al frame 12 della sequenza «Renata» codificata a circa 1 bit/pel, con mascheramento degli errori random dal tipo 1 al tipo 8 (vedere tabella 5). Si confronta l'efficienza del mascheramento degli errori con e senza l'utilizzo dei vettori movimento.

Fig. 1B — Mascheramento degli errori ottenuto sostituendo il blocco errato con quello coposizionato appartenente al field od al frame precedente.

Fig. 2B — Mascheramento degli errori ottenuto sostituendo il quadriblocco errato con quello preso dal field precedente (blocco errato codificato in modo intra-field od inter-field), oppure ricavato dal frame precedente lungo la direzione indicata dal vettore movimento (blocco errato codificato in modo inter-frame).

APPENDICE C

Le fotografie riportate sono relative al frame 12 della sequenza «Renata», con compensazione del movimento e con fattore di trasmissione FT = 72 corrispondente ad un bit-rate medio di 1 bit/pel, in presenza di errore random ed a burst, con probabilità di occorrenza $P(e) = 10^{-3}$ e con differenti strategie di mascheramento degli errori.

Fig. 1C — *Errore random*. Nessun mascheramento degli errori.

Fig. 2C — *Errore random*. Mascheramento degli errori di tipo da 1 a 6 (tabella 5) effettuato a livello di quadriblocco.

Fig. 3C — *Errore random*. Mascheramento degli errori di tipo da 1 a 7 (tabella 5) effettuato a livello di quadriblocco.

Fig. 4C — *Errore random.* Mascheramento degli errori di tipo da 1 a 8 (tabella 5) effettuato a livello di 8 quadriblocchi.

Fig. 5C — *Errore a burst.* Nessun mascheramento degli errori.

Fig. 6C — *Errore a burst.* Mascheramento degli errori di tipo da 1 a 6 (tabella 5) effettuato a livello di quadriblocco.

Fig. 7C — *Errore a burst.* Mascheramento degli errori di tipi da 1 a 7 (tabella 5) effettuato a livello di quadriblocco.

Fig. 8C — *Errore a burst.* Mascheramento degli errori di tipo da 1 a 8 (tabella 5) effettuato a livello di 8 quadriblocchi.

BIBLIOGRAFIA

1. - M. BARBERO, M. STROPPIANA: *Codifica del segnale televisivo numerico: uso della trasformata coseno discreta*, «Elettronica e Telecomunicazioni», n. 1 - 1989.
2. - M. BARBERO, R. DEL PERO, P. GIROMINI: *Codifica del segnale televisivo numerico: struttura di trama per un sistema basato su DCT*, «Elettronica e Telecomunicazioni», n. 1 - 1990.
3. - N. ZENONI: *A hybrid DCT codec: studies to maintain a good picture quality in the presence of transmissions errors*, «Assises des Jeunes Chercheurs», Tokyo, 8-12 June '92.
4. - M. BARBERO, R. DEL PERO, M. STROPPIANA, N. ZENONI: *Digital Distribution of TV and HDTV signals: Strategies to Maintain Good Picture Quality in the Presence of Transmission Errors* «3rd International Workshop of Digital Signal Processing Techniques Applied to Space Communications», 23-25 september 1992, ESTEC, Nordwijk, The Netherlands.
5. - CCIR Racc. 723
6. - ETS 300 174: *Network applications: digital coding of component television signals for contribution quality applications in the range 34-45 Mbit/s.*
7. - M. BARBERO, R. BELLORA, M. STROPPIANA: *Codifica del segnale televisivo numerico: codici a lunghezza variabile applicati ai coefficienti DCT, «Elettronica e Telecomunicazioni», n. 1 - 1990.*

“Elettronica e Telecomunicazioni”,
nata nel 1952 come “Elettronica e Televisione Italiana”, è una rivista quadrimestrale di Rai Eri realizzata dal Centro Ricerche e Innovazione Tecnologica della Rai, sul cui sito è disponibile gratuitamente dal 2001.

Il Centro Ricerche e Innovazione Tecnologica (CRIT) della Rai nasce a Torino nel 1930 come “Laboratorio Ricerche” e dal 1960 ha sede in Corso Giambone 68. Successivamente assume la denominazione “Centro Ricerche” e, dall’ottobre 1999, quella attuale.
L’attività del Centro è coordinata dalla Direzione Strategie Tecnologiche.
Alla nascita, tra i suoi obiettivi ha la progettazione e realizzazione di impianti ed apparati di nuova concezione, non reperibili sul mercato. I profondi cambiamenti nello scenario delle telecomunicazioni hanno stimolato la trasformazione del Centro.
Ha ricevuto riconoscimenti a livello internazionale per i contribuiti forniti alle attività di studio e normalizzazione dei sistemi per la codifica dei segnali audio e video in forma digitale, allo sviluppo delle tecniche di compressione dei segnali attualmente alla base dei sistemi di trasmissione e registrazione dei segnali video, alla definizione degli standard di diffusione e trasmissione DVB.
Il Centro contribuisce all’evoluzione delle tecnologie relative al sistema radiotelevisivo e multimediale e supporta il Gruppo nelle scelte di indirizzo tecnologico e nella fase di sperimentazione e introduzione in esercizio di nuovi prodotti e sistemi. E’ attivo in numerosi progetti finanziati in ambito europeo e nazionale e collabora con Università e Industrie per l’attività di ricerca, per la definizione dei nuovi standard e lo sviluppo dei nuovi servizi.

Rai Radiotelevisione S.p.A.
Centro Ricerche e Innovazione Tecnologica
Corso E. Giambone, 68 - I 10135 Torino
www.crit.rai.it